Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients

Ralph G. Andrzejak,[1] Kaspar Schindler,[2] and Christian Rummel[3]

[1]*Universitat Pompeu Fabra, Department of Information and Communication Technologies, E-08018 Barcelona, Spain*

[2]*qEEG group, Department of Neurology, Inselspital, Bern University Hospital, University of Bern, CH-3010 Bern, Switzerland*

[3]*Support Center for Advanced Neuroimaging, University Institute for Diagnostic and Interventional Neuroradiology, Inselspital, Bern University Hospital, University of Bern, CH-3010 Bern, Switzerland*

(Received 3 August 2012; revised manuscript received 4 September 2012; published 12 October 2012)

To derive tests for randomness, nonlinear-independence, and stationarity, we combine surrogates with a nonlinear prediction error, a nonlinear interdependence measure, and linear variability measures, respectively. We apply these tests to intracranial electroencephalographic recordings (EEG) from patients suffering from pharmacoresistant focal-onset epilepsy. These recordings had been performed prior to and independent from our study as part of the epilepsy diagnostics. The clinical purpose of these recordings was to delineate the brain areas to be surgically removed in each individual patient in order to achieve seizure control. This allowed us to define two distinct sets of signals: One set of signals recorded from brain areas where the first ictal EEG signal changes were detected as judged by expert visual inspection ("focal signals") and one set of signals recorded from brain areas that were not involved at seizure onset ("nonfocal signals"). We find more rejections for both the randomness and the nonlinear-independence test for focal versus nonfocal signals. In contrast more rejections of the stationarity test are found for nonfocal signals. Furthermore, while for nonfocal signals the rejection of the stationarity test increases the rejection probability of the randomness and nonlinear-independence test substantially, we find a much weaker influence for the focal signals. In consequence, the contrast between the focal and nonfocal signals obtained from the randomness and nonlinear-independence test is further enhanced when we exclude signals for which the stationarity test is rejected. To study the dependence between the randomness and nonlinear-independence test we include only focal signals for which the stationarity test is not rejected. We show that the rejection of these two tests correlates across signals. The rejection of either test is, however, neither necessary nor sufficient for the rejection of the other test. Thus, our results suggest that EEG signals from epileptogenic brain areas are less random, more nonlinear-dependent, and more stationary compared to signals recorded from nonepileptogenic brain areas. We provide the data, source code, and detailed results in the public domain.

DOI: 10.1103/PhysRevE.86.046206 PACS number(s): 05.45.Tp, 05.45.Xt, 87.19.le, 87.19.lm

I. INTRODUCTION

Nonlinear signal analysis comprises a wide variety of measures that allow one to extract different characteristic features of dynamical systems underlying experimental signals [1]. Applications to signals measured from the brain, for example, contribute to our understanding of brain functions and malfunctions and thereby help to advance cognitive neuroscience and neurology [2,3]. In particular electroencephalographic (EEG) recordings from epilepsy patients attract researchers' interest. In some epilepsy patients the diagnostics requires to record the EEG directly from the surface of the brain or from deeper brain structures. The clinical purpose of these intracranial recordings is to localize the brain areas where seizures start and to assess whether the patient can benefit from the neurosurgical resection of these parts of the brain. From the physics' point of view, these recordings reveal intriguing dynamics not only during acute epileptic seizures but also during the seizure-free interval. Therefore, such intracranial recordings from epilepsy patients are a prominent and challenging field of applications for nonlinear signal analysis [4]. There is growing evidence that this interdisciplinary analysis can contribute valuable diagnostic information about the localization of the epileptic focus even from the seizure-free interval [5–9]. This is highly important, because each seizure is a potentially health impairing event and epileptologists strive to minimize the number of seizures that have to be recorded for diagnostic epilepsy surgery evaluation.

Univariate nonlinear measures estimate features such as the dimensionality, predictability, or entropy of individual dynamics X from single signals x. Pairs of signals x and y measured simultaneously from two dynamics X and Y are analyzed using bivariate nonlinear measures to detect interactions between the dynamics. However, both univariate and bivariate nonlinear measures have an important limitation. While they are sensitive to characteristic features of nonlinear dynamics they often lack specificity. The problem is that nonlinear measures are strongly influenced by linear correlations of the signals, and arbitrary degrees of linear auto- and cross-correlation can, for example, be obtained for linear stochastic processes. For EEG signals cross-correlation can reflect the underlying dynamics but can likewise be caused by volume conduction or the reference montage.

This lack of specificity can be addressed with the concept of surrogates, first origins of which can be found in studies of electrocardiographic [10] and electroencephalographic [11] signals. Originally devised as a test for nonlinearity in univariate dynamics [12], the concept of surrogates has become a very versatile tool in signal analysis. Different types of surrogates can be generated for univariate [12–14] as well as for bivariate

or multivariate signals [15–19]. While classically surrogates are combined with nonlinear measures, their combination with linear measures [9,20–25] is equally straightforward and useful. Exploiting these different variations of the concept of surrogates one can test a wide variety of null hypotheses about the dynamics underlying experimental signals.

Surrogate signals are generated by randomizing the original signals. This randomization is constrained such that selected properties of the original signals are preserved. In particular, the composition of the maintained properties can be adapted to the different null hypothesis [26]. A discriminating statistics, which has to be sensitive to at least one signal property that is not consistent with the null hypothesis, is calculated for both the original signal and the surrogates. If the result for the original deviates from the distribution of the values obtained from the surrogates, the null hypothesis is rejected.

As a concrete example suppose that we have calculated some low value of a nonlinear prediction error (Ref. [1] and references therein) from an experimental signal. We wonder whether our result indicates nonlinear deterministic structure of the underlying dynamics or simply reflects the signal's linear autocorrelations. Since we cannot answer this question based on the nonlinear prediction error alone, we generate a set of surrogate signals of the original signal. The surrogates are constrained to have the same linear autocorrelations as the original signal but to be otherwise random. We calculate the nonlinear prediction error also for the surrogate signals and test whether the original result is inside or outside the surrogates' result range. If it is outside, the linear autocorrelations are not sufficient to explain the low value of the nonlinear prediction error, and the dynamics is not consistent with a stationary linear stochastic process. This does not prove that the dynamics is nonlinear deterministic. The surrogates' null hypothesis is composed of various assumptions, and the violation of any of these assumptions renders it incorrect. In consequence different alternative models remain for our dynamics. To further narrow down these alternatives we can proceed and analyze our signal with other types of surrogates in combination with other nonlinear or also linear measures [27].

There is growing evidence that the additional information gained by surrogates can be decisive for a successful characterization of EEG recordings of epilepsy patients [5–9,11,20,23–25,27–32]. Investigating intracranial periseizure EEG recordings in rats, Pijn *et al*. [11] were the first to combine a nonlinear measure and surrogates to analyze recordings from the brain. Casdagli *et al*. [5] and Andrzejak *et al*. [6] combined different univariate nonlinear measures and surrogates to analyze intracranial EEG recordings from patients with unilateral medial temporal lobe epilepsy. In both studies rejections of the null hypothesis of a linear stochastic process were prevalent in signals measured in the seizure-generating brain area. Extending earlier work of Ref. [21], Rummel *et al*. [23] combined the linear cross-correlation with univariate surrogates to extract the "genuine cross-correlation" between individual channel pairs. In comparison to the raw cross-correlation, the genuine cross-correlation better assessed spatiotemporal interaction patterns from an intracranial periseizure EEG recording of an epilepsy patient. In particular, the genuine cross-correlation revealed the strong involvement of focal signals in these interactions. In Ref. [27] this concept was extended by including mutual information as nonlinear bivariate interrelation measure as well as multivariate surrogates. Rejections of the null hypothesis of linear interrelations were found predominantly for the EEG recorded in epileptogenic brain areas as well as during epileptic seizures.

In a study of continuous EEG recordings of the seizure-free interval in 29 patients with medial temporal lobe epilepsy Andrzejak *et al*. [7] tested a variety of univariate measures. They showed that combinations of nonlinear measures with univariate surrogates allowed to localize the epileptic focus in a high percentage of cases. Importantly, this approach clearly outperformed the use of nonlinear measures without surrogates as well as linear measures. Recently Andrzejak *et al*. [8] showed that these findings carry over to bivariate signal analysis. Based on the same EEG recordings they showed that a combination of a nonlinear interdependence measure with bivariate surrogates excels nonlinear interdependence measures without surrogates as well as the linear cross correlation in localizing the epileptic focus. In both studies more rejections of the surrogate null hypotheses were obtained for the EEG recorded in the focal as compared to the nonfocal EEG. The authors concluded that focal EEG signals were distinct from nonfocal EEG signals in that they are less consistent with an underlying linear stochastic process and rather reflect some properties of a coupled nonlinear deterministic dynamics [8].

However, as indicated above, the rejection of a surrogates' null hypothesis always leaves one with different alternative interpretations. As already indicated by Andrzejak *et al*. [7,8], apart from nonlinear determinism their results could be explained by the nonstationarity of the EEG. The assumption of stationarity is included in the null hypotheses of the univariate and bivariate surrogates used in these studies. Therefore, if the focal EEG was more nonstationary than the nonfocal EEG and if this nonstationarity favored the rejection of the null hypotheses tested with univariate or bivariate nonlinear measures that were used in these studies, this could explain the increased rejection rate found for the focal EEG.

Another important open question is whether univariate and bivariate surrogate analysis can provide much nonredundant information about the dynamics underlying the EEG. In the first place univariate and bivariate nonlinear signal analysis measures extract in general independent aspects of the dynamics. For example, we can have two low-dimensional nonlinear deterministic dynamics. This structure would be detected by any sensitive univariate nonlinear measure. However, that does not imply whether or not there is dependence between these two dynamics to be detected by bivariate measures. The dynamics may or may not be coupled. On the other hand, any nonrandom structure in either dynamics violates both the univariate and bivariate null hypothesis. Whether the hypotheses are actually rejected due to this nonrandomness in turn depends on the type of the univariate and bivariate measures used to test the hypothesis.

Therefore, our first objective is to assess the influence of nonstationarity on a univariate randomness test and a bivariate nonlinear-independence test in application to focal and nonfocal EEG signals. We base the randomness test on a nonlinear prediction error and univariate surrogates, and the nonlinear-independence test on a nonlinear interdependence measure and bivariate surrogates. As stationarity test we use

a combination of linear fluctuation measures with univariate surrogates. Our second objective is to study whether the rejections of the univariate randomness test and the bivariate nonlinear-independence test correlate across signals. To minimize the influence of nonstationarity and the differences between the focal and nonfocal dynamics, we restrict this part of the study to focal signals for which the stationarity test was not rejected.

II. METHODS

A flowchart providing an overview of the different steps of analysis described in this section is provided in Appendix C.

A. Presurgical epilepsy diagnostics

We included intracranial EEG recordings from five epilepsy patients. These recordings were performed prior to and independently from our study as part of the epilepsy diagnostics in these patients. All five patients had longstanding pharmacoresistant temporal lobe epilepsy and were candidates for epilepsy surgery. Noninvasive studies had not allowed for unequivocal localization of the brain areas from which seizures originated ("seizure onset zone"), and all patients underwent long-term intracranial EEG recordings at the Department of Neurology of the University of Bern. Multichannel EEG signals were recorded with intracranial strip and depth electrodes all manufactured by AD-TECH (Racine, WI, USA). An extracranial reference electrode placed between 10/20 positions Fz and Pz was used. EEG signals were either sampled at 512 or 1024 Hz, depending on whether they were recorded with more or less than 64 channels. Based on these intracranial EEG recordings the brain areas where seizures started could be localized for all five patients. In addition, these areas were found in parts of the brain that could be surgically resected without the danger of neurological deficits that would be unacceptable for the patients. All five patients had good surgical outcome. Three patients attained complete seizure freedom, and two patients only had auras but no other seizures following surgery, corresponding to class 1 and 2 according to the "International League Against Epilepsy" classification of surgical outcome [33]. Retrospective EEG data analysis has been approved by the ethics committee of the Kanton of Bern. In addition, all patients gave written informed consent that their data from long-term EEG might be used for research purposes.

B. Preprocessing of EEG signals

All EEG signals were digitally band-pass filtered between 0.5 and 150 Hz using a fourth-order Butterworth filter. Forward and backward filtering was used in order to minimize phase distortions. Those EEG signals that had been recorded with a sampling rate of 1024 Hz were down-sampled to 512 Hz prior to further analysis. EEG signals were then re-referenced against the median of all the channels free of permanent artifacts as judged by visual inspection. There is no reference that can be considered "best" on general grounds. Rummel *et al*. [34] investigated the impact of six common EEG references on bi- and multivariate correlation measures for the example of scalp montages. The global average clearly introduces correlation, however, in a controlled way. The median reference was not investigated in this publication. However, as compared to the mean it has the additional advantage that the rank of the correlation matrix remains full. In consequence, the amount of artificially introduced correlation is even smaller than for the global average reference.

C. Composition of sets of EEG signals

As "focal EEG channels" we defined all those channels that detected first ictal EEG signal changes as judged by visual inspection by at least two neurologists who are also board-certified electroencephalographers. KS was always one of the experts. Though visual analysis is not a fully objective approach, joint-analysis with fellow neurologists allows reducing subjective interpretation. Furthermore, visual EEG analysis is currently still the most important technique for clinical decision making (see Sec. II A). All other channels included in the recordings were classified as "nonfocal EEG channels." We randomly selected 3750 pairs of simultaneously recorded signals x and y from the pool of all signals measured at focal EEG channels. For that purpose, we at first divided the recordings into time windows of 20 seconds, corresponding to 10 240 samples. Recordings of seizure activity and three hours after the last seizure were excluded. For each individual signal pair we then randomly selected one of the five patients, one of this patient's focal EEG channels (for the signal x), one of this channel's neighboring focal channels (for the signal y), and one time window included in this patient's recordings. This random sample was drawn without replacement and using a uniform random number generator. Before being included into the database, the signal pair was visually inspected. In case, it contained prominent measurement artifacts, the signal pair was discarded. Moderate contaminations by power-line noise at 50 Hz, however, were not used as exclusion criterion. No clinical selection criteria such as the presence or absence of epileptiform activity were applied. Finally, the focal EEG signal pairs were stored in the order in which they were drawn. Their origin (patient, channel, window) was not stored. In the same way we randomly selected 3750 pairs of nonfocal signals measured at nonfocal EEG channels. Exemplary pairs of signals are shown in Fig. 1 along with the outcomes of the different tests described below (see also Sec. V).

D. Surrogate signals

For the different hypotheses tests described below we used univariate [13] and bivariate surrogate signals [16]. *Univariate* surrogates are generated from univariate signals. Accordingly, given the pair of signals x and y, univariate surrogate signals are generated by randomizing both individual signals separately. The randomization is constrained such that the surrogates have the same autocorrelation and amplitude distribution as the original signals. Any potential nonlinear deterministic structure or nonstationary features of the original signals are destroyed. Pairs of *bivariate* surrogate signals are generated by randomizing the pair of original signals jointly. Like in the univariate case, the surrogates have the same autocorrelation and amplitude distribution as the original

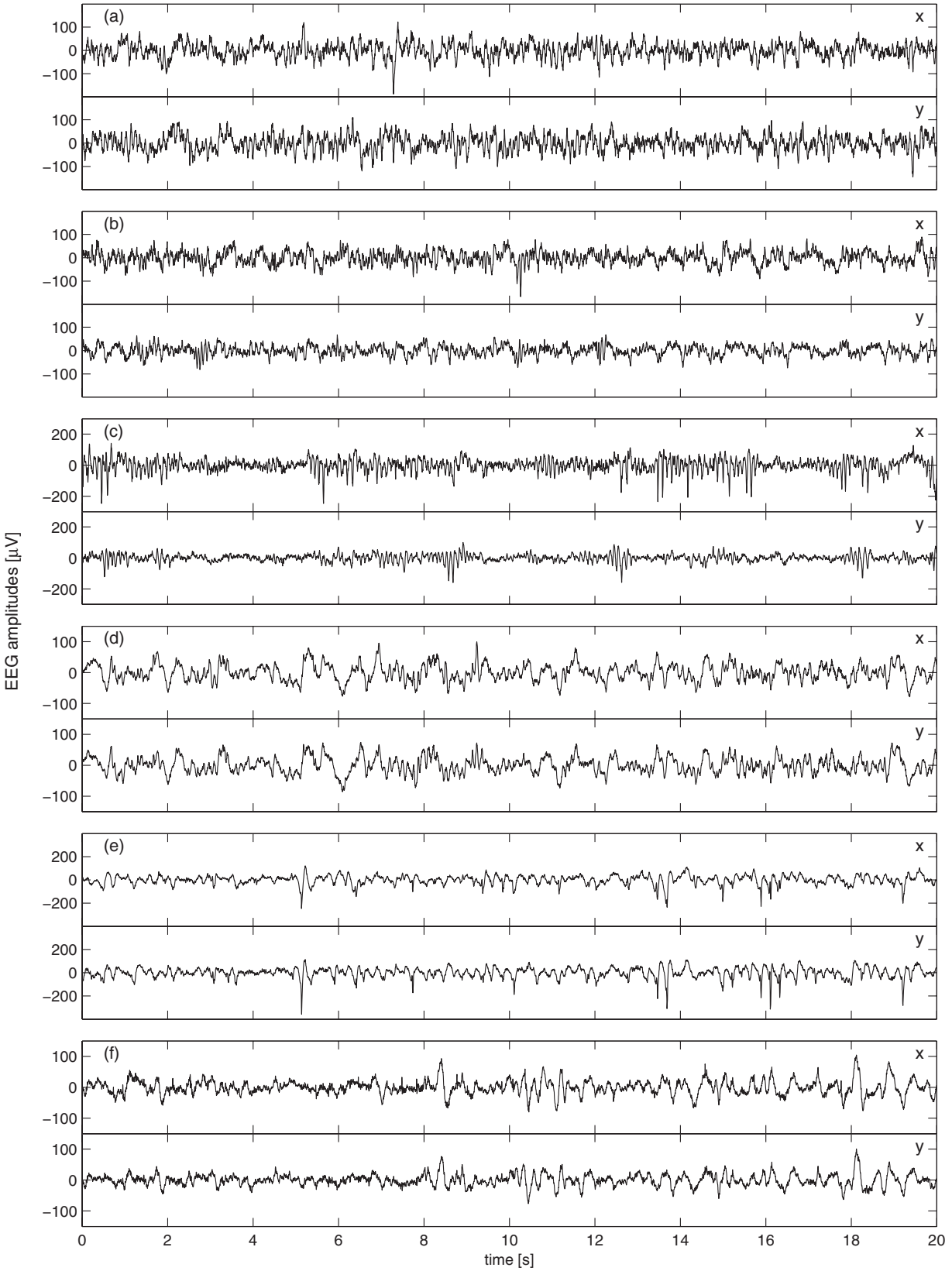

FIG. 1. Exemplary pairs of nonfocal signal pairs (a–c) and focal signal pairs (d–f). The test outcomes for these signals are (a) $\mathcal{S}_0, \mathcal{U}_0^{\mathrm{X}}, \mathcal{U}_0^{\mathrm{Y}}, \mathcal{B}_0$; (b) $\mathcal{S}_0, \mathcal{U}_0^{\mathrm{X}}, \mathcal{U}_1^{\mathrm{Y}}, \mathcal{B}_1$; (c) $\mathcal{S}_1, \mathcal{U}_1^{\mathrm{X}}, \mathcal{U}_1^{\mathrm{Y}}, \mathcal{B}_1$; (d) $\mathcal{S}_0, \mathcal{U}_0^{\mathrm{X}}, \mathcal{U}_0^{\mathrm{Y}}, \mathcal{B}_0$; (e) $\mathcal{S}_0, \mathcal{U}_1^{\mathrm{X}}, \mathcal{U}_1^{\mathrm{Y}}, \mathcal{B}_1$; (f) $\mathcal{S}_1, \mathcal{U}_0^{\mathrm{X}}, \mathcal{U}_0^{\mathrm{Y}}, \mathcal{B}_0$.

signals. Beyond that the pair of surrogates also has the same cross-correlation as the original pair of signals. Again any potential nonlinear deterministic structure or nonstationary features of the original signals are destroyed. Furthermore, any signature of nonlinear interdependence between x and y is removed.

The null hypothesis tested by univariate surrogates $\mathcal{H}_{0,\mathrm{univ}}$ is that *the dynamics is a stationary linear stochastic correlated Gaussian process. The measurement function by which the signal was derived from the dynamics is invertible but potentially nonlinear. The autocorrelation, mean, and variance of the underlying Gaussian process are such that the measurement results in the autocorrelation, and amplitude distribution of the observed time series*. The null hypothesis of the bivariate surrogate signals $\mathcal{H}_{0,\mathrm{biv}}$ is a generalization of the univariate version: *The dynamics is a stationary bivariate linear stochastic correlated Gaussian process. The measurement functions by which the pair of signals was derived from the dynamics are invertible but potentially nonlinear. The autocorrelation, cross-correlation, mean, and variance of the underlying Gaussian process are such that the measurement results in the autocorrelation, cross-correlations and amplitude distribution of the observed time series*.

We generated the surrogates signals using an iterative algorithm proposed for the univariate case in Ref. [13] and generalized to the bivariate case in Ref. [16]. Detailed descriptions of these algorithms can be found in these original references or in Ref. [35] (see also Sec. V).

E. Randomness tests: $\mathcal{U}_0^{\mathrm{X}}$ and $\mathcal{U}_0^{\mathrm{Y}}$ versus $\mathcal{U}_1^{\mathrm{X}}$ and $\mathcal{U}_1^{\mathrm{Y}}$

The randomness test is based on a nonlinear prediction error N (Ref. [1] and references therein) and univariate surrogates. It is carried out separately for the signals x and y. The nonlinear prediction error N aims at a distinction between stochastic and deterministic dynamics. For this purpose, it quantifies the degree to which similar present states are mapped to similar future states of the dynamics. The dynamics is at first reconstructed from the signal using delay coordinates, and similarity of states is assessed by spatial proximity in the reconstructed state space. For a detailed description of the algorithm used to calculate the measure N we refer to Ref. [6] (see also Sec. V).

For periodic dynamics the nonlinear prediction error N takes values of zero. The other extreme is uncorrelated white noise for which N values around one are obtained. When calculated for a signal from a deterministic dynamics and for a stochastic signal that has the same autocorrelation like the deterministic signal, lower nonlinear prediction errors are typically obtained for the deterministic signal. However, stochastic but strongly autocorrelated signals can result in lower N values than weakly autocorrelated deterministic signals. Hence, only a combination of the nonlinear prediction error and surrogates can serve as randomness test.

To calculate N we low-pass filtered the signals with a cut-off frequency of 40 Hz (eighth-order Butterworth filter) and subsequently down-sampled the signals by a factor of four, resulting in a sampling time of 7.8 ms. Subsequently, the state space was reconstructed using an embedding dimension m and time delay τ. The parameters of the nonlinear prediction error are the number of nearest neighbors k, the prediction horizon H, and the Theiler correction window W. In Secs. III A–III B, we show detailed results obtained for $m = 8$, $\tau = 4$ sampling times, $k = 5$, $H = 4$ sampling times, and $W = 19$ sampling times. The parameter dependence of the results is summarized in Appendix A.

We calculated the nonlinear prediction error N for the signal x and for 19 univariate surrogate signals generated from it. We rejected the randomness test if the result for the signal was lower than the minimal result across all 19 surrogate signals. Accordingly, the randomness test has a significance level of $\alpha = 0.05$. The outcome of a rejection of the randomness test for x is denoted by $\mathcal{U}_1^{\mathrm{X}}$. If the N value of at least one surrogate was lower than the one for the original signal, the randomness test for x was not rejected. This outcome is denoted by $\mathcal{U}_0^{\mathrm{X}}$. All steps were carried out analogously but independently for y resulting in the two possible outcomes $\mathcal{U}_1^{\mathrm{Y}}$ and $\mathcal{U}_0^{\mathrm{Y}}$.

F. Nonlinear-independence test: $\mathcal{B}_0$ versus $\mathcal{B}_1$

For the nonlinear-independence test we use a nonlinear interdependence measure L and bivariate surrogates. The nonlinear interdependence measure L aims at a characterization of couplings between two dynamics X and Y from the analysis of a pair of signals measured from them [36]. Like a number of related approaches [1,18,37–42], the measure L quantifies the probability with which similar states of one dynamics are mapped to similar states of the other dynamics. Like for the nonlinear prediction error, the dynamics are at first reconstructed from the signals x and y using delay coordinates, and similarity of states is assessed by spatial proximity in the respective reconstructed state spaces. Due to an unbiased normalization and the use of a rank-based statistics, the measure L offers a higher sensitivity and specificity for directional couplings than a number of previous approaches [36].

In its elementary form this approach provides two directional measures: $L(X|Y)$ and $L(Y|X)$. Both directional measures take values distributed around zero for signals of independent realizations of independent dynamics. For weak couplings from X to Y the measure $L(X|Y)$ increases while $L(Y|X)$ remains close to zero. Analogously, weak couplings in the other direction can be detected from an increase of $L(Y|X)$ and $L(X|Y)$ being close to zero. Once the coupling is strong enough to induce a synchronized motion of X and Y, both measures attain high values, and the upper bound of $L(Y|X) = L(X|Y) = 1$ is reached for identical synchronization. Accordingly, the difference $L(X|Y) - L(Y|X)$ can be used to characterize the direction of the coupling, as long as the coupling is not strong enough to induce synchronization. Following Ref. [8] we restrict ourselves to the characterization of the overall strength of the nonlinear dependence and use the nondirectional symmetrized version of the measure:

$$L = \frac{L(X|Y) + L(Y|X)}{2}. \tag{1}$$

For a detailed description of the algorithm used to calculate L we refer to Ref. [36] (see also Sec. V).

To calculate the nonlinear interdependence measure L we used the same filtering and down-sampling as for the nonlinear prediction error. We also used the same parameter values for the embedding dimension m, time delay τ, number nearest neighbors k, and the Theiler correction window W. No prediction horizon H is needed for the measure L. Accordingly, results of the nonlinear-independence test shown in Secs. III A–III B are obtained for $m = 8$, $\tau = 4$ sampling

times, $k = 5$, and $W = 19$ sampling times. The parameter dependence of our results is summarized in Appendix A.

Apart from a coupling between nonlinear deterministic dynamics, cross-correlations resulting from superpositions of independent dynamics can result in high values of L. Accordingly, only a combination of the nonlinear interdependence measure L with bivariate surrogates can be used as nonlinear-independence test. Therefore, we calculated the measure L for the original pair of signals x and y and for 19 pairs of surrogate signals. We rejected the nonlinear-independence test if the result for the pair of original signals was higher than the maximal result across all 19 pairs of surrogate signals, again corresponding to a test with a significance level of $\alpha = 0.05$. The outcome of a rejection of the nonlinear-independence test is denoted by $\mathcal{B}_1$. If the result of at least one pair of surrogates exceeded the results for the pair of original signals, the nonlinear-independence test was not rejected. This outcome is denoted by $\mathcal{B}_0$. We use the term "nonlinear-interdependence test" instead of just "independence test" since the null hypothesis includes linear cross-correlation between the two signals.

G. Stationarity test: $\mathcal{S}_0$ versus $\mathcal{S}_1$

For the stationarity test we combine an amplitude-stationarity test and frequency-stationarity already used in [6], and a correlation-stationarity test. Since these tests are nonstandard and somewhat *ad hoc*, we provide the full formulas here. Given the pair of signals x and y of 10 240 samples each, the following steps are carried out. At first the signal are both normalized to zero mean and unit variance. Then both are divided into 16 nonoverlapping subsegments of length 640 samples: $x_{i,j}$ and $y_{i,j}$ for $i = 1, \ldots, 16$ and $j = 1, \ldots, 640$. For each subsegment we calculate the average absolute deviation across the amplitudes

$$A_{x_i} = \frac{1}{640} \sum_{j=1}^{640} |x_{i,j} - \overline{x_{i,j}}|, \tag{2}$$

where the overbar denotes the mean across the samples of segment i. A_{y_i} is calculated analogously. Furthermore the mean frequency is determined from

$$F_{x_i} = \frac{\sum_{k=1}^{320} f_k S_{x_i}(f_k)}{\sum_{k=1}^{320} S_{x_i}(f_k)}, \tag{3}$$

where $S_{x_i}(f_k)$ denotes the amplitude of the Fourier transform of the subsegment i at frequency f_k. F_{y_i} is calculated analogously. The equal-time cross-correlation coefficient between corresponding segments of x and y is calculated using

$$C_i = \frac{1}{640} \sum_{j=1}^{640} \frac{(x_{i,j} - \overline{x_{i,j}})(y_{i,j} - \overline{y_{i,j}})}{\sigma(x_{i,j})\sigma(y_{i,j})}, \tag{4}$$

where $\sigma(\cdot)$ denotes the standard deviation across the samples of segment i.

The fluctuation of these quantities across subsegments is quantified using the average deviations

$$R(A_x) = \frac{1}{16} \sum_{i=1}^{16} |A_{x_i} - \langle A_{x_i} \rangle|, \tag{5}$$

where the brackets $\langle \cdot \rangle$ denote the mean across the segments, $R(A_y), R(F_x), R(F_y), R(C)$ analogously. $R(A_x)$ and $R(F_x)$ are calculated for the signal x and 99 univariate surrogates generated from x, $R(A_y)$ and $R(F_y)$ in turn for y and 99 univariate surrogates of y. The quantity $R(C)$ is calculated for the pair of signals x and y and 99 bivariate surrogates generated from them.

The stationarity test is designed to be very strict. To reject it, it is sufficient that one of the five values obtained for the original signals [$R(A_x)$, $R(A_y)$, $R(F_x)$, $R(F_y)$, and $R(C)$] is outside the range of its surrogates. This outcome is denoted by $\mathcal{S}_1$. We use a higher number of 99 surrogates to avoid that the chance level of this combined test exceeds 5%. If we regard the amplitude-stationarity test, frequency-stationarity, and correlation-stationarity test as independent, the significance level of the stationarity test is $1 - 0.99^5 = 0.049$. The outcome that the stationarity test is not rejected is denoted by $\mathcal{S}_0$.

H. Outcomes: Counts, probability estimates, and terminology

We denote the counts of the different test outcomes across signal pairs by $c(\cdot)$. Since we have two univariate randomness tests, one bivariate nonlinear-independence test, and one stationarity test, we can have up to four arguments for the counts. Whenever we refer to marginal counts, we suppress all arguments over which the sum was taken, e.g., $c(\mathcal{U}_0^X\mathcal{U}_0^Y\mathcal{B}_1\mathcal{S}_1) + c(\mathcal{U}_1^X\mathcal{U}_0^Y\mathcal{B}_1\mathcal{S}_1) + c(\mathcal{U}_0^X\mathcal{U}_0^Y\mathcal{B}_0\mathcal{S}_1) + c(\mathcal{U}_1^X\mathcal{U}_0^Y\mathcal{B}_0\mathcal{S}_1) = c(\mathcal{U}_0^X\mathcal{U}_0^Y\mathcal{S}_1) + c(\mathcal{U}_1^X\mathcal{U}_0^Y\mathcal{S}_1) = c(\mathcal{U}_0^Y\mathcal{S}_1)$. In consequence, when we refer to the sum of all counts marginalized over all tests, we use no argument for c.

From the outcome counts $c(\cdot)$ we define estimates of probabilities $p(\cdot)$ and conditioned probabilities $p(\cdot|\cdot)$. For the sake of brevity we often drop the term *estimate*. We, however, have to keep in mind that these values derived from outcome counts are subject to fluctuations caused by the finite sample size. Confidence intervals of these probability estimates are derived in Appendix B.

Like the counts of outcomes, the derived probabilities can have up to four arguments. Among the arguments can be none, one, or both outcomes of the two univariate randomness tests, e.g., $p(\mathcal{B}_1)$, $p(\mathcal{U}_1^X|\mathcal{B}_1)$, $p(\mathcal{U}_1^Y|\mathcal{B}_1\mathcal{U}_0^X)$, respectively. Recall, however, that our database is constructed to be symmetric with regard to the univariate properties of X and Y. Accordingly all probabilities that have at least one univariate randomness test outcome among the arguments are paired. The definition of one paired probability can be transformed to its counterpart by exchanging the symbols X and Y. The values of both probabilities are identical, e.g., $p(\mathcal{U}_1^X|\mathcal{B}_1) = p(\mathcal{U}_1^Y|\mathcal{B}_1)$, or $p(\mathcal{U}_1^Y|\mathcal{B}_1\mathcal{U}_0^X) = p(\mathcal{U}_1^X|\mathcal{B}_1\mathcal{U}_0^Y)$. The values of our estimates of these probabilities coincide, except for fluctuations caused by the finite sample. To avoid redundancies, we therefore show only one of the paired probabilities, namely, the one where the outcome of X occurs alone or first, e.g., $p(\mathcal{U}_1^X|\mathcal{B}_1)$ or $p(\mathcal{U}_1^X|\mathcal{B}_1\mathcal{U}_0^Y)$. Importantly, that does not mean that we pool across the two univariate randomness tests to improve the statistics. We cannot pool across these two tests since they cannot assumed to be based on independent samples. (In Secs. III B4–III B5 we study the actual dependence between the two univariate randomness test outcomes.)

To compare probabilities we use the relative difference:

$$D(p_1,p_2)=\frac{p_1-p_2}{p_1+p_2}. \tag{6}$$

Evidently this quantity is bounded by $[-1,1]$, positive and negative values of $D(p_1,p_2)$ are obtained for $p_1>p_2$ and $p_1<p_2$, respectively. Confidence intervals for this quantity are derived in Appendix B.

As indicated above and further discussed in Sec. IV, the outcomes $\mathcal{B}_1$ and $\mathcal{B}_0$ can indicate but cannot prove nonlinear-dependence and nonlinear-independence, respectively. The same holds for $\mathcal{U}_1^{\mathrm{X}}$, $\mathcal{U}_1^{\mathrm{Y}}$ versus $\mathcal{U}_0^{\mathrm{X}}$, $\mathcal{U}_0^{\mathrm{Y}}$ with regard to nonrandomness and randomness, as well as for $\mathcal{S}_1$ and $\mathcal{S}_0$ with regard to nonstationarity and stationarity. This has to be kept in mind when we use these terms in the following.

III. RESULTS

A. Focal versus nonfocal signals: Randomness, nonlinear independence, and stationarity

In this section we compare the rejection counts of the randomness tests and the nonlinear-independence test for the focal and nonfocal signals. We use the index f and n, respectively, to distinguish between these signal classes. In formulas the index a is used as place holder for both: $a=\{f,n\}$.

1. More focal nonrandomness and nonlinear dependence

Estimates of the rejection probabilities for the randomness and nonlinear-independence test are given by, respectively,

$$p_a\left(\mathcal{U}_1^{\mathrm{X}}\right)=\frac{c_a\left(\mathcal{U}_1^{\mathrm{X}}\right)}{c_a}, \tag{7}$$

$$p_a(\mathcal{B}_1)=\frac{c_a(\mathcal{B}_1)}{c_a}. \tag{8}$$

Figure 2(a) shows that all estimates are clearly above the chance level of 5% and that rejection probabilities for the nonlinear-independence test are higher than the ones for the randomness test. More importantly, for both tests we obtain higher rejection probabilities for the focal signals as for the nonfocal signals.

2. More nonfocal nonstationarity

The estimates of the rejection probabilities for the stationarity test are

$$p_a(\mathcal{S}_1)=\frac{c_a(\mathcal{S}_1)}{c_a}. \tag{9}$$

Figure 3 shows that in contrast to the randomness and nonlinear-independence test, more rejections of the stationarity test are found for the nonfocal signals.

3. Impact of nonstationarity stronger for nonfocal signals

Does the stationarity test outcome have an influence on the rejection probabilities of the randomness and nonlinear-independence tests? To address this question we contrast the overall probability $p_a(\mathcal{U}_1^{\mathrm{X}})$ with the conditioned probabilities:

$$p_a\left(\mathcal{U}_1^{\mathrm{X}}\middle|\mathcal{S}_1\right)=\frac{c_a\left(\mathcal{U}_1^{\mathrm{X}}\mathcal{S}_1\right)}{c_a(\mathcal{S}_1)}, \tag{10}$$

$$p_a\left(\mathcal{U}_1^{\mathrm{X}}\middle|\mathcal{S}_0\right)=\frac{c_a\left(\mathcal{U}_1^{\mathrm{X}}\mathcal{S}_0\right)}{c_a(\mathcal{S}_0)}. \tag{11}$$

We furthermore compare $p_a(\mathcal{B}_1)$ to the corresponding conditioned probabilities, which are defined analogously to Eqs. (10) and (11). For the nonfocal signals a rejection of the stationarity test increases the rejection probabilities for both the randomness and the nonlinear-independence test substantially [Fig. 2(c)]. In contrast, for the focal signals a rejection of the stationarity test has almost no impact on the rejection probabilities for the randomness test. It does increases the rejection probability for the nonlinear-independence test, however, to a lesser extent than for the nonfocal signals, as evidenced by the D values displayed in Figs. 2(c)–2(d).

Keep in mind that the results shown in Figs. 2(c)–2(d) are not a consequence of the ones shown in Fig. 3. The different stationarity test rejection probabilities found for the nonfocal versus nonfocal signals do not imply what happens with the rejection probabilities of the other tests, given that the stationarity was rejected or was not rejected.

4. Increased contrast between focal and nonfocal signals for stationary signals

Results of the previous section imply that the contrast between the focal and nonfocal signals, found by the randomness test and the nonlinear-independence test [Fig. 2(a)], changes if we exclude signals for which the stationarity test was rejected. Specifically, regarding Figs. 2(c)–2(d), we expect that this contrast increases. Indeed, Fig. 2(b) shows that the D values obtained for the quantities conditioned on that the stationarity test is not rejected are substantially higher than those for the unconditioned counterparts shown in Fig. 2(a).

B. Dependence between nonlinear independence and randomness tests: Stationary, focal signals only

In this section we study the dependence between the randomness and the nonlinear-independence test outcomes. Results of Sec. III A show that the focal and nonfocal signals have different statistical properties with regard to these two tests. Therefore, to isolate the dependence between the tests, we restrict the following analysis to focal signals. To simplify the notation we drop the index f. Furthermore, Sec. III A3 shows that the outcome of the stationarity test influences the rejection probabilities of the randomness and nonlinear-independence tests. Even though this influence is weaker for the focal signals, we restrict our analysis further by excluding those signals, for which the stationarity test was rejected. Thereby we aim to minimize the influence of nonstationarity. After this reduction of our data base we are left with 2000 focal signal pairs. In consequence, all subsequent expressions are conditioned on $\mathcal{S}_0$. We suppress this conditioning argument to simplify the notation. Note that it is a pure coincidence that we have a round number of exactly 2000 focal signal pairs for which the stationarity test was not rejected.

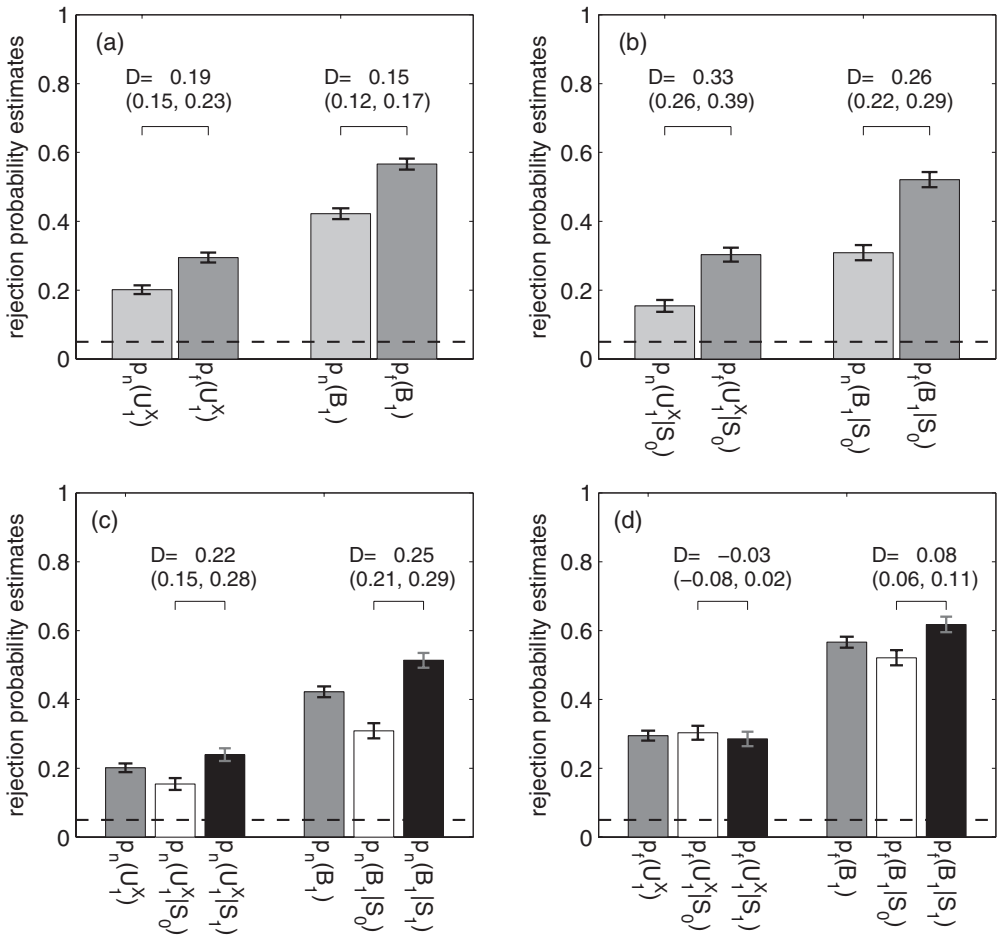

FIG. 2. (a) More focal nonrandomness and nonindependence. Comparison of rejection probabilities for nonfocal (light gray) and focal signals (dark gray) for the randomness test (left) and nonlinear-independence test (right). (b) Increased contrast between focal and nonfocal signals for stationary signals. Composed like panel (a), but here the analysis is restricted to signals for which the stationarity test was not rejected. (c) Impact of nonstationarity for nonfocal signals. Comparison of probabilities of the outcome that the randomness test is rejected (left) and the outcome that the nonlinear-independence test is rejected (right). Probabilities are conditioned on that the stationarity test is rejected (black) or not rejected (white). For gray bars the stationarity test outcome is marginalized out. (d) Impact of nonstationarity for focal signals. Composed like panel (c), but here for focal signals. Graphical elements: bars: rejection probability estimates; error bars: 95% confidence intervals; vertical lines: significance level of the tests; D values: determined according to Eq. (6) for the pair of probabilities above which they are displayed, along with 95% confidence intervals (see Appendix B for confidence intervals).

1. More nonlinear dependence for nonrandom signals

Does the rejection probability of the nonlinear-independence test depend on the outcomes of the randomness tests? To address this question we compare the conditional probability

$$p\left(\mathcal{B}_1|\mathcal{U}_0^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}}\right)=\frac{c\left(\mathcal{U}_0^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}}\mathcal{B}_1\right)}{c\left(\mathcal{U}_0^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}}\right)} \tag{12}$$

to the analogously defined conditional probabilities $p(\mathcal{B}_1|\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}})$, and $p(\mathcal{B}_1|\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_1^{\mathrm{Y}})$. Figure 4(a) shows that already the rejection of one of the two randomness tests increases the rejection probability of the nonlinear-independence test. The rejection of both randomness tests further increases this probability.

We find however also that even if no randomness test is rejected, there is a substantial probability to reject the nonlinear-independence test: $p(\mathcal{B}_1|\mathcal{U}_0^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}})$ is very far from the chance level. Accordingly, the rejection of the randomness test is not necessary for the nonlinear-independence test rejection. Furthermore, even if both randomness tests are rejected, there is a substantial probability not to reject the bivariate test: $p(\mathcal{B}_1|\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_1^{\mathrm{Y}})$ remains clearly below one. Hence, the rejection of the randomness test is not sufficient for the rejection of the nonlinear-independence test.

2. Nonrandomness does not cause false positive detections of nonlinear dependence

We showed in the previous subsection that the rejection of the randomness tests is not sufficient for the rejection of the nonlinear-independence test. Nonetheless, results displayed in Fig. 4 could suggest that nonrandomness in X or Y, and even more so in X and Y, favors the rejection of the nonlinear-independence test, even if X and Y are independent.

To test this possible conjecture we generated independent signals by breaking the pairing between the x and y signals.

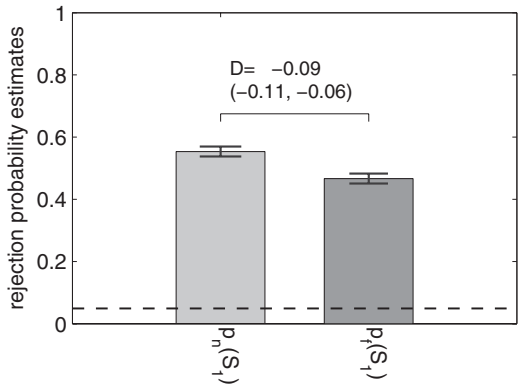

FIG. 3. More nonfocal nonstationarity. Comparison of the rejection probabilities of the stationarity test for the nonfocal (light gray) and focal signals (dark gray). All graphical elements like in Fig. 2.

Instead of pairing the signal $x^{(l)}$ with its simultaneously measured counterpart $y^{(l)}$, we paired it with some other $y^{(k\neq l)}$ signal, which was selected without replacement from the pool of all remaining y signals. (Accordingly, the upper indexes l,k refer to the number of the signal within all 2000 stationary focal signals. They do not refer to samples within individual signals. The signal samples were not shuffled.) We then repeated the analysis for these shuffled signal pairs. In particular, we again excluded all signals for which the stationarity test was rejected. The shuffled signal pairs have the same statistical properties with regard to the randomness test as the nonshuffled pairs. However, the shuffling destroys any potential dependence. Figure 4(b) shows that for these shuffled signal pairs, the rejection probability of $\mathcal{H}_{0,\mathrm{biv}}$ is very close to the chance level of 5%. In all cases, the confidence intervals include this significance level. In particular, no evident dependence on the outcome of the randomness test is found. Hence, the nonlinear-independence test is not rejected for independent X and Y.

3. More nonrandomness for nonlinear dependent signals

To study in which way the rejection probabilities of the randomness test depend on the outcomes of the nonlinear-independence test, we use the conditional probability

$$p\left(\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_1^{\mathrm{Y}}|\mathcal{B}_1\right)=\frac{c\left(\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_1^{\mathrm{Y}}\mathcal{B}_1\right)}{c(\mathcal{B}_1)} \tag{13}$$

as well as the analogously defined conditional probabilities $p(\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_1^{\mathrm{Y}}|\mathcal{B}_0)$, $p(\mathcal{U}_0^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}}|\mathcal{B}_1)$, $p(\mathcal{U}_0^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}}|\mathcal{B}_0)$, $p(\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}}|\mathcal{B}_1)$, and $p(\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}}|\mathcal{B}_0)$. We find that the rejection of the nonlinear-independence test increases the probability to reject the randomness test (Fig. 5). In particular, the probability to reject both random tests jointly is increased substantially, while the probability to reject just one randomness test is only moderately increased.

Furthermore, we find that even if the nonlinear-independence test is not rejected, there is a high probability to still reject at least one of the randomness tests, $p(\mathcal{U}_0^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}}|\mathcal{B}_0)$ stays clearly below one. Hence, the rejection of the nonlinear-independence test is not necessary for the randomness test rejection. Furthermore, even if the nonlinear-independence test is rejected, there remains a substantial probability $p(\mathcal{U}_0^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}}|\mathcal{B}_1)$ to accept both randomness tests. Therefore, the rejection of the nonlinear-independence test is not sufficient for the rejection of the randomness test.

4. $\mathcal{U}_1^{\mathrm{X}}$ and $\mathcal{U}_1^{\mathrm{Y}}$ are dependent

Assume that $\mathcal{U}_1^{\mathrm{X}}$ and $\mathcal{U}_1^{\mathrm{Y}}$ were independent events with a certain probability q. In this case we expect $p(\mathcal{U}_0^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}})\approx(1-q)^2$, $p(\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}})\approx p(\mathcal{U}_0^{\mathrm{X}}\mathcal{U}_1^{\mathrm{Y}})\approx q(1-q)$, and $p(\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_1^{\mathrm{Y}})\approx q^2$. However, results displayed in Fig. 5 suggest that $p(\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_1^{\mathrm{Y}})$ is too high compared to $p(\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}})$ to be consistent with this assumption of $\mathcal{U}_1^{\mathrm{X}}$ and $\mathcal{U}_1^{\mathrm{Y}}$ being independent events. Direct evidence for their dependence can be obtained from

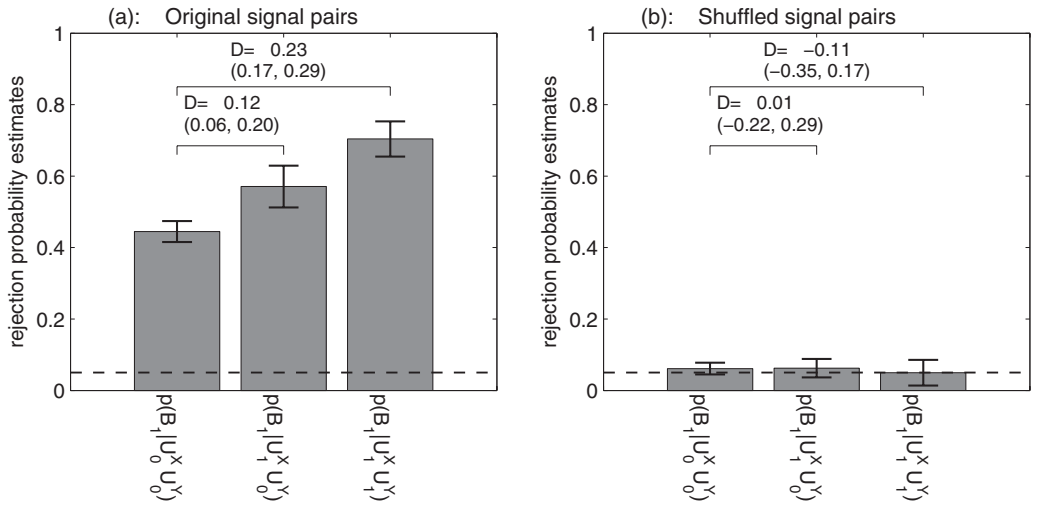

FIG. 4. (a) More nonlinear dependence for nonrandom signals. Comparison of the rejection probabilities for the nonlinear-independence test conditioned on that no randomness test is rejected (left), one randomness test is rejected (middle) and both randomness tests are rejected (right). (b) Nonrandomness does not cause false positive detections of dependence. Composed like panel (a), but here for shuffled signal pairs. All graphical elements like in Fig. 2.

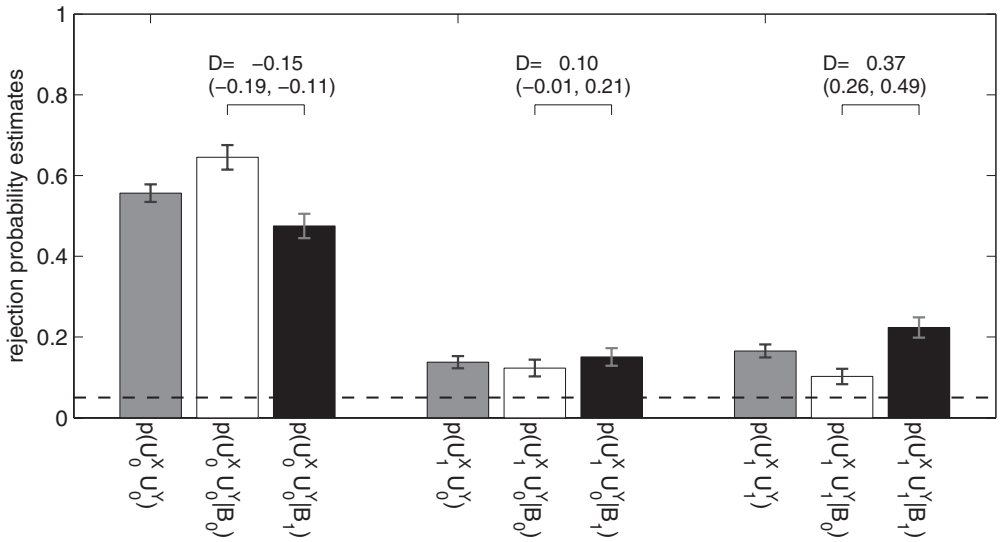

FIG. 5. More nonrandomness for nonlinear dependent signals. Comparison of probabilities of the joint outcome that no randomness test is rejected (left), one randomness test is rejected (middle), and both randomness tests are rejected (right). We show these probabilities conditioned on that the nonlinear-independence test is rejected (black) or not rejected (white). For gray bars the nonlinear-independence test outcome is marginalized out. All graphical elements like in Fig. 2.

the conditional probabilities:

$$p(\mathcal{U}_1^X|\mathcal{U}_1^Y) = \frac{c(\mathcal{U}_1^X\mathcal{U}_1^Y)}{c(\mathcal{U}_1^Y)} \tag{14}$$

and the analogously defined $p(\mathcal{U}_1^X|\mathcal{U}_0^Y)$. If $\mathcal{U}_1^X$ and $\mathcal{U}_1^Y$ were independent events, we expect $p(\mathcal{U}_1^X|\mathcal{U}_1^Y) \approx p(\mathcal{U}_1^X|\mathcal{U}_0^Y)$ and accordingly $D[p(\mathcal{U}_1^X|\mathcal{U}_1^Y), p(\mathcal{U}_1^X|\mathcal{U}_0^Y)] \approx 0$. However, this relation clearly does not hold [Fig. 6(a)], providing direct evidence for the dependence of $\mathcal{U}_1^X$ and $\mathcal{U}_1^Y$.

5. The dependence of $\mathcal{U}_1^X$ and $\mathcal{U}_1^Y$ depends on $\mathcal{B}_1$

To quantify the degree of dependence between $\mathcal{U}_1^X$ and $\mathcal{U}_1^Y$ we contrast the conditional probability $p(\mathcal{U}_1^X|\mathcal{U}_1^Y)$ with the probabilities that are additionally conditioned on the nonlinear-independence test outcome:

$$p(\mathcal{U}_1^X|\mathcal{U}_1^Y\mathcal{B}_1) = \frac{c(\mathcal{U}_1^X\mathcal{U}_1^Y\mathcal{B}_1)}{c(\mathcal{U}_1^Y\mathcal{B}_1)} \tag{15}$$

and $p(\mathcal{U}_1^X|\mathcal{U}_1^Y\mathcal{B}_0)$ defined analogously. Figure 6(b) shows that the dependence between the randomness test outcomes across X and Y is increased if the nonlinear-independence test is rejected.

IV. DISCUSSION

For both the randomness test and the nonlinear-independence test we obtain more rejections for the focal signals as compared to the nonfocal signals (Sec. III A1). These findings are consistent with previous findings of Refs. [5–7] and [8,23,27], respectively. An important potential confounding variable in such results derived from surrogate tests is nonstationarity. Since stationarity is included in $\mathcal{H}_{0,\mathrm{univ}}$ and $\mathcal{H}_{0,\mathrm{biv}}$, nonstationarity can cause the rejection of the randomness and the nonlinear-independence test. While the potential effect of nonstationarity is often discussed or stationarity of the signals is sometimes even used as an inclusion criterion for the signals, results of Secs. III A2–III A4 for the first time assess the impact of nonstationarity on the rejection probabilities of a randomness and nonlinear-independence test applied to EEG time series. Regarding the high rejection rates of the stationarity test (Sec. III A2), we should recall that we designed this test to be very strict. To reject it, it is sufficient to reject the frequency-stationarity test or the amplitude-stationarity test for x or y or the correlation-stationarity test for the pair x and y. Nonetheless, rejection rates around 50% reflect that on the time scale of the analysis window our EEG signals show strong fluctuations of the frequencies, amplitude magnitudes and correlations. These fluctuations exceed in a large fraction of signals those found for the surrogates leading to the high rejection rates of the stationarity test.

It is important, however, to keep in mind that also our stationarity test is based on a null hypothesis test. Accordingly, for the principal reasons outlined in the introduction, its rejection cannot prove that the underlying dynamics is nonstationary. Consider for example the following simple stationary stochastic process. The first variable u_i counts the number of events emitted by a homogeneous Poisson process up to time index i. The second variable v_i is set to zero if u_i is even and set to one if u_i is odd. The third variable is defined as $w_{i+1} = 0.5 w_i u_i + \xi_i$, where ξ_i is Gaussian white noise. Evidently, depending on the ratio between the Poisson process rate and the window length the stationarity test can be rejected when applied to w. Another example is a signal from the deterministic Lorenz dynamics. This dynamics involves two time scales, fast oscillations within individual wings and slow switchings between the wings of the butterfly attractor. In the case these irregular switches occur seldom with regard to the window length used for the stationarity test, the test can be rejected. In both examples the dynamics are stationary, but the stationarity test can be rejected depending on the ratio between the long time scales of the dynamics and the window length

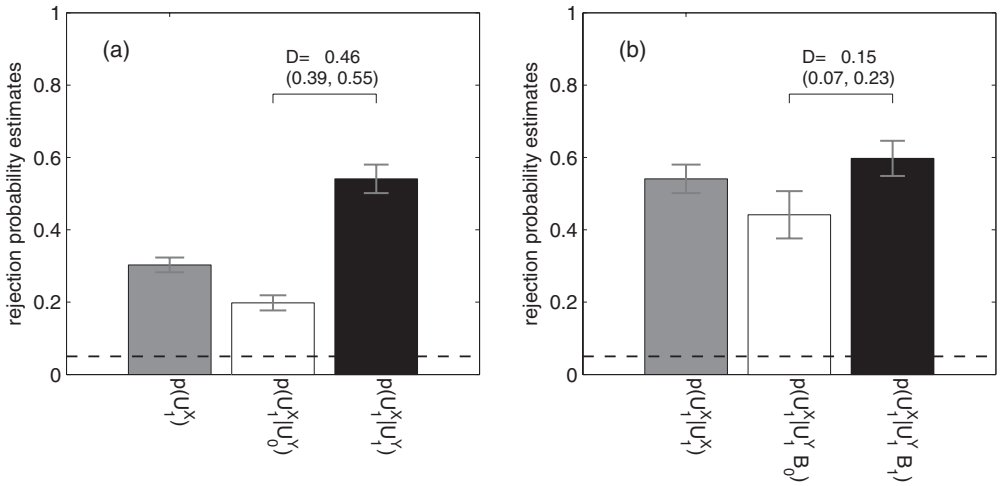

FIG. 6. (a) $\mathcal{U}_1^X$ and $\mathcal{U}_1^Y$ are dependent. Comparison of probabilities of the outcome that the randomness test is rejected for X. We show these probabilities conditioned on that the randomness test for Y is rejected (black) or not rejected (white). For gray bars the randomness test outcome for Y is marginalized out. (b) The outcome $\mathcal{B}_1$ enhances the dependence between $\mathcal{U}_1^X$ and $\mathcal{U}_1^Y$. Comparison of the probabilities that the randomness test for X is rejected given that it has been rejected for Y. We show these probabilities conditioned on that the nonlinear-independence test is rejected (black) or not rejected (white). For the gray bar the nonlinear-independence test outcome is marginalized out. Accordingly this bar is a replica of the black bar in panel (a). All graphical elements like in Fig. 2.

used for the stationarity test. However, in fact these rejections are not false but true positive rejections. The stationarity test is a test of $\mathcal{H}_{0,\mathrm{univ}}$ based on a measure of nonstationarity. In both examples, the dynamics are not consistent with $\mathcal{H}_{0,\mathrm{univ}}$, both processes are nonlinear. The Lorenz dynamics is moreover deterministic. Accordingly, the rejection of the test is correct in both examples. It would be incorrect, however, to interpret this rejection as proof for the nonstationarity of the dynamics. On the other hand, whatever causes the fluctuations of the mean frequency, mean amplitudes, and linear correlations of the EEG signals, their effect is that the signals seem nonstationary on the time scale of the analysis window. After all that is what counts when the EEG signals are used as input for the randomness and nonlinear-independence test. In that sense we interpret results of Sec. III A2 to indicate nonstationarity.

Due to the intermittent occurrence of epileptiform activity during the seizure-free interval one might expect that focal signals are less stationary than nonfocal signals. However, in consistency with an early study of Wang and Wieser [43], we in fact find more nonstationarity for nonfocal signals than for the focal signals. Another important finding is that the impact of the rejection of the stationarity test on the rejection probabilities of the other tests is distinct for focal versus nonfocal signals (Sec. III A3). For nonfocal signals the rejection of the stationarity test increases the rejection probabilities for the randomness and the nonlinear-independence test substantially. For focal signals the rejection of the stationarity test increases the rejection probabilities for the nonlinear-independence test, while for the randomness test even a slight decrease of the rejection probability is found. Overall the impact is clearly weaker for the focal versus nonfocal signals. As a consequence, the contrast between the focal and nonfocal signals, found by the randomness test and the nonlinear-independence test, enhances when we exclude signals for which the stationarity test is rejected (Sec. III A4). Certainly these findings deserve further investigation. How can one characterize these different ratios and types of nonstationarity found for the focal and nonfocal signals (cf. Ref. [44])? Do these findings depend on different recording locations (temporal lobe, occipital lobe, parietal lobe, frontal lobe, hippocampal formation)? Do these findings depend on the time scale used for the stationarity test? These questions are beyond the scope of the present study and are left for future studies, to which everyone is invited to contribute (see Sec. V).

In summary of Sec. III A, nonstationarity is indeed a confounding variable for the randomness and nonlinear-independence test. However, once we control for it, the significance of the main result, namely, the contrast between the focal and nonfocal signals obtained by these tests, is actually further enhanced. So, while we challenged the conclusion of Ref. [8] in the introduction, our results further support it. Accordingly, in keeping with Ref. [8] we interpret the results of Sec. III A to reflect an increased level of synchronization of groups of neurons induced by epilepsy during the seizure-free interval. This increased synchronicity can have different manifestations on different spatial scales of neuronal organization. For univariate focal signals recorded by individual contacts it can cause the rejection of the randomness test. For bivariate focal signals recorded by pairs of contacts it can cause the rejection of the nonlinear-independence test. Across these spatial scales this focal synchronicity can result in an EEG that is less consistent with a linear stochastic process and more consistent with a coupled nonlinear deterministic dynamics. These results for the seizure-free interval complement studies which provide evidence for nonlinear deterministic dynamics in epileptic seizures (e.g., Refs. [11,35,45]). We do not imply that epilepsy induces a transition from a pure linear stochastic process to a pure coupled nonlinear deterministic dynamics. Rather our interpretation is that the EEG reflects the superposition of these two types of dynamics and that epilepsy strengthens the coupled nonlinear deterministic fraction. This strengthening and the complexity of the coupled nonlinear deterministic

fraction will vary across different signal pairs, so that the randomness and nonlinear-independence tests are rejected only for a portion of the signals.

It is important to note, however, that also the present results can not yield a proof for an underlying nonlinear deterministic dynamics. As stated in the introduction, the reason is that the null hypotheses $\mathcal{H}_{0,\mathrm{univ}}$ and $\mathcal{H}_{0,\mathrm{biv}}$ are composed of a number of distinct assumptions, and the violation of any of these assumptions renders these hypotheses wrong. In consequence the rejection of the randomness test cannot prove that the dynamics is nonlinear deterministic, and the joint rejection of the randomness test and the nonlinear-independence test cannot provide a proof for a coupled nonlinear deterministic dynamics. We provided strong evidence against that our results are caused by the nonstationarity of the dynamics. Thereby we provided evidence against one, but only one, alternative to the conclusion of an underlying nonlinear deterministic dynamics. Many alternatives remain, such as a nonlinear or non-Gaussian stochastic process. To test for these is left for future work.

In Sec. III B we only included focal signals for which the stationarity test was not rejected in order to assess the correlation between the randomness and nonlinear-independence test outcomes. We find that the rejections of the randomness test and nonlinear-independence test are positively correlated across signals (Secs. III B1 and III B3). In particular, the rejection of the nonlinear-independence test increases the probability that the randomness test is rejected for both x and y. Analogously the outcomes of only one or both randomness tests being rejected both increase the rejection probability of the nonlinear-independence test. In particular, the increase is stronger when both randomness tests are rejected. These findings provide further support for the conclusion drawn from Sec. III A that the focal EEG contains a nonlinear deterministic part. This violates both $\mathcal{H}_{0,\mathrm{univ}}$ and $\mathcal{H}_{0,\mathrm{biv}}$, and both null hypotheses are wrong and should be jointly rejected. On the other hand, we find that the rejection of the randomness test is neither necessary nor sufficient for the rejection of the nonlinear-independence test. Analogously the rejection of the nonlinear-independence test is neither necessary nor sufficient for the rejection of the randomness test. This indicates that the features induced by epilepsy become sometimes predominantly evident from the structure of individual signals and sometimes rather manifests themselves in the nonlinear interdependence between signals.

Extending results for mathematical model systems [36,46] and time-shifted surrogates of EEG signals [8], results of Sec. III B2 confirm the excellent specificity of the measure L. They show that the rejection probability of the nonlinear-independence test obtained for independent signals is consistent with chance level. This holds regardless of whether the randomness test is rejected for none, one, or both univariate signals. Let us suppose that the rejection of the randomness tests is caused since X and Y are a pair of nonlinear deterministic but independent dynamics. Despite being independent, the individual dynamics' deterministic structure violates $\mathcal{H}_{0,\mathrm{biv}}$. However, due to the high specificity of the particular nonlinear interdependence measure L it takes values distributed around zero for independent dynamics, regardless of possible deterministic structure in the univariate signals. Accordingly, for a pair of nonlinear deterministic but independent dynamics, values distributed around zero are obtained for the original signals and for the surrogates. In consequence, the nonlinear-independence test is not rejected. Thanks to this high specificity an nonlinear-independence test based on this measure and surrogates can indeed provide specific information about the interrelation of the two dynamics and is not influenced by the structure of the individual dynamics.

The dependence between the randomness test outcomes across x and y (Sec. III B4), which is further increased if the nonlinear-independence test is rejected (Sec. III B5), is plausible. We only included signals that where measured from neighboring contacts. EEG signals measured from neighboring contacts are frequently but not necessarily strongly correlated. Hence, these are not independent measurements, and the randomness test derived from these signals cannot be expected to be independent. The increase of this dependence upon rejection of the nonlinear-independence test is analogous to the finding that a nonlinear-independence test rejection increases the probability that the randomness test is rejected for both x and y (Secs. III B1 and III B3).

Our results show that the randomness test and nonlinear-independence test extract nonredundant information from focal EEG signals. In consequence, both tests can contribute different aspects to a thorough characterization of EEG signals measured from epilepsy patients. Both tests can help to distinguish focal and nonfocal signals. Importantly, for both tests this contrast is enhanced when we include only signals for which the stationarity test was not rejected. Future work shall study how combinations of the tests used here and potential further tests can be optimized to localize brain areas where seizures originate without the necessity to observe actual seizure activity.

Recordings from the brain have always been a challenging application for nonlinear signal analysis. Problems encountered in the study of brain signals often promoted the improvement of existing nonlinear measures, the development of new measures, or even new concepts such as surrogates. Our results provide further evidence that the benefit of this interdisciplinary field is mutual. The application of nonlinear signal analysis can provide valuable clinical information.

V. Sharing of data, source code, and results

The data, source code, and detailed results are provided in Ref. [47]. The data comprises all EEG signals analyzed here. The source code includes the algorithms for the calculation of the measures N and L, the generation of the univariate and bivariate surrogates, as well as for the stationarity test. The outcomes of the stationarity test, randomness test, and nonlinear-independence test are given for each individual pair of EEG signals. In this way, the interested reader can inspect the data and the corresponding results without the need to rerun the analysis. The detailed results also allow one to determine whatever joined or conditioned rejection probability which can be of interest but was not included here. This resource of data, source code, and results can be extended in all aspects. We therefore invite the scientific community to contribute further results obtained with other measures for the data studied here, or further data to be evaluated with the tests used here.

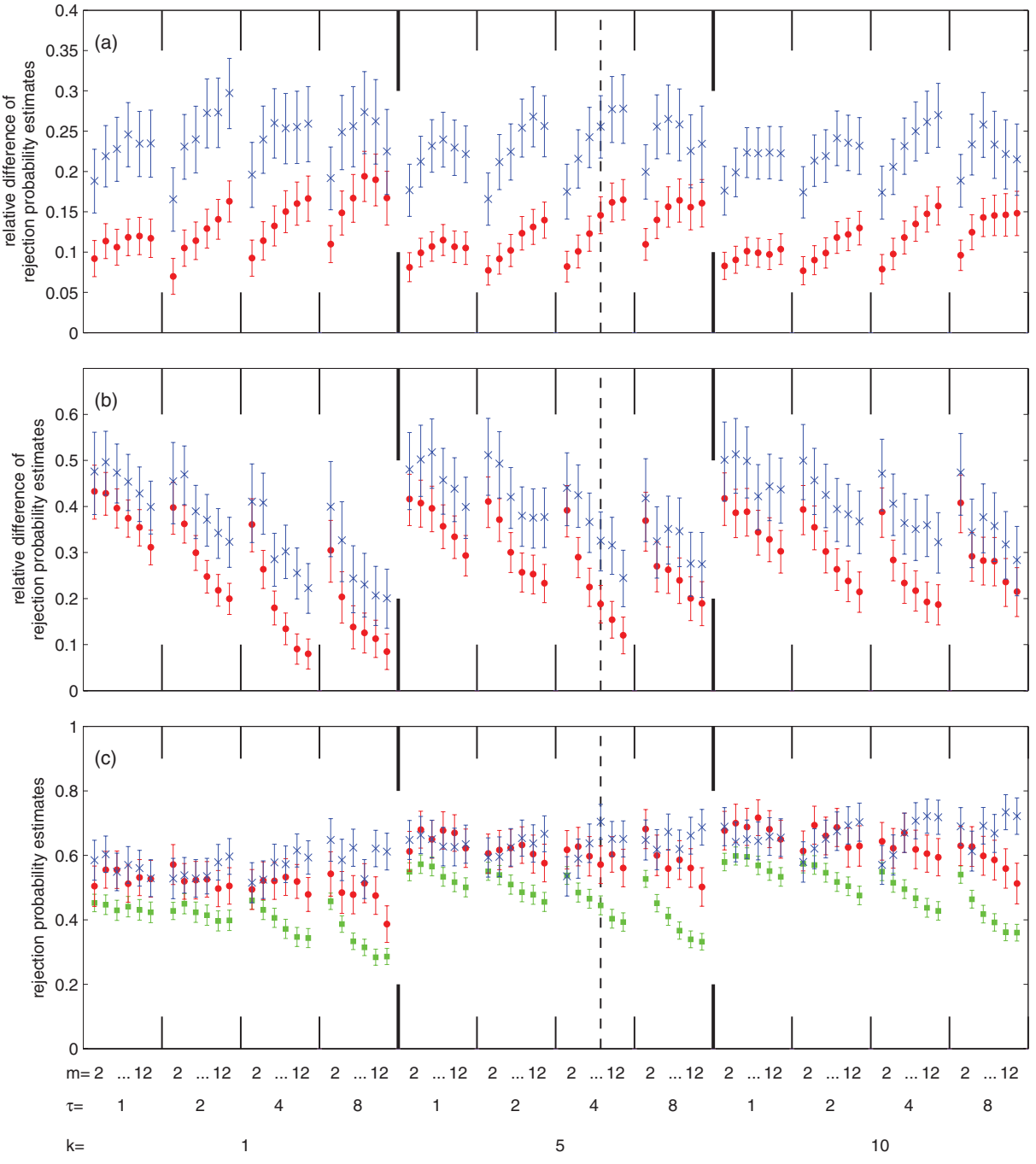

FIG. 7. (Color online) (a) $D[p_f(\mathcal{B}_1), p_n(\mathcal{B}_1)]$ (red circles) and $D[p_f(\mathcal{B}_1|\mathcal{S}_0), p_n(\mathcal{B}_1|\mathcal{S}_0)]$ (blue crosses) in dependence on m, τ, and k. (b) $D[p_f(\mathcal{U}_1^{\mathrm{X}}), p_n(\mathcal{U}_1^{\mathrm{X}})]$ (red circles) and $D[p_f(\mathcal{U}_1^{\mathrm{X}}|\mathcal{S}_0), p_n(\mathcal{U}_1^{\mathrm{X}}|\mathcal{S}_0)]$ (blue crosses) in dependence on m, τ and k for a fixed $H = 4$ sampling times. (c) Dependence of $p(\mathcal{B}_1|\mathcal{U}_0^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}})$ (green squares, always below red circles), $p(\mathcal{B}_1|\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_0^{\mathrm{Y}})$ (red circles), and $p(\mathcal{B}_1|\mathcal{U}_1^{\mathrm{X}}\mathcal{U}_1^{\mathrm{Y}})$ (blue crosses) in dependence on m, τ, and k for a fixed $H = 4$ sampling times. Note the different scaling of the y axes of the different panels. In all panels the parameter setting used in the main text is indicated by the vertical line.

ACKNOWLEDGMENTS

R.G.A. acknowledges grant FIS-2010-18204 of the Spanish Ministry of Education and Science. K.S. and C.R. are grateful for support by the Swiss National Science Foundation (projects nos. SNF 320030-122010 and 33CM30-124089).

APPENDIX A: PARAMETER DEPENDENCE

For the nonlinear prediction error and the nonlinear interdependence measure, underlying the randomness and nonlinear-independence test, respectively, we varied the parameters of the embedding dimension within $m = [2,4,6,8,10,12]$ and the time delay within $\tau = [1,2,4,8]$ sampling times. The number

of nearest neighbors was varied within $k = [1,5,10]$. The nonlinear prediction error has the prediction horizon H as additional parameter which we varied within $H = [1,2,4,8,16,32]$ sampling times. The Theiler correction window was fixed to $W = 19$ sampling times. For the stationarity test the only parameter is the number of subsegments. We did not vary this parameter but fixed it to 16 according to Ref. [6].

In general, the rejection probabilities of the randomness test and the nonlinear-independence test depend on the different parameters. However, our main findings, derived from comparisons of these probabilities, show only a weak parameter dependence. Deviations are mostly found for small values of m and τ. In this context we should note that the temporal distance between the first and last entry in delay coordinate vectors is given by $(m-1)\tau$, and a sufficient value of this so-called embedding window is necessary to properly reconstruct complex dynamics.

An exhaustive description of the dependence of all results on all parameters would be far too lengthy. Therefore we restrict ourselves to three representative examples. Figure 7(a) shows the dependence of $D[p_f(\mathcal{B}_1), p_n(\mathcal{B}_1)]$ and $D[p_f(\mathcal{B}_1|\mathcal{S}_0), p_n(\mathcal{B}_1|\mathcal{S}_0)]$ on m, τ, and k. We see that across all parameter values:

$$D[p_f(\mathcal{B}_1|\mathcal{S}_0), p_n(\mathcal{B}_1|\mathcal{S}_0)] > D[p_f(\mathcal{B}_1), p_n(\mathcal{B}_1)] > 0. \quad \text{(A1)}$$

This means that more nonlinear-independence test rejections are always found for focal versus nonfocal signals. Furthermore, this contrast is always increased when the analysis is restricted to signals for which the stationarity test was not rejected. Both $D[p_f(\mathcal{B}_1), p_n(\mathcal{B}_1)]$ and $D[p_f(\mathcal{B}_1|\mathcal{S}_0), p_n(\mathcal{B}_1|\mathcal{S}_0)]$ tend to decrease for increasing k. With regard to increasing m and τ these quantities either increase or show a dependence of $\cap$ shape.

For the randomness test we have the additional parameter of the prediction horizon H. Figure 7(b) shows the dependence of $D[p_f(\mathcal{U}_1^X), p_n(\mathcal{U}_1^X)]$ and $D[p_f(\mathcal{U}_1^X|\mathcal{S}_0), p_n(\mathcal{U}_1^X|\mathcal{S}_0)]$ on m, τ, and k for a fixed $H = 4$ sampling times. We see that for this value of the prediction horizon across all values of the other parameters:

$$D\left[p_f\left(\mathcal{U}_1^X|\mathcal{S}_0\right), p_n\left(\mathcal{U}_1^X|\mathcal{S}_0\right)\right] > D\left[p_f\left(\mathcal{U}_1^X\right), p_n\left(\mathcal{U}_1^X\right)\right] > 0 \quad \text{(A2)}$$

This means that, at $H = 4$, for the randomness test more rejections are found for the focal signals across all parameter combinations. This is in agreement with findings for the nonlinear-independence test. Again this contrast is increased across all parameter values when the analysis is restricted to signals for which the stationarity test was not rejected. Both $D[p_f(\mathcal{U}_1^X|\mathcal{S}_0), p_n(\mathcal{U}_1^X|\mathcal{S}_0)]$ and $D[p_f(\mathcal{U}_1^X), p_n(\mathcal{U}_1^X)]$ tend to decrease for increasing m, increasing τ, and decreasing k. Similar results are found for smaller and higher values of the prediction horizon H. Only for $H = 1, m = 4, \tau = 1, k = 5$ and $H = 1, m = 4, \tau = 1, k = 10$ we found negative $D[p_f(\mathcal{U}_1^X), p_n(\mathcal{U}_1^X)]$ values. For all other 430 parameter combinations we found positive $D[p_f(\mathcal{U}_1^X), p_n(\mathcal{U}_1^X)]$ values. $D[p_f(\mathcal{U}_1^X|\mathcal{S}_0), p_n(\mathcal{U}_1^X|\mathcal{S}_0)]$ is positive for all 432 parameter combinations and higher than $D[p_f(\mathcal{U}_1^X), p_n(\mathcal{U}_1^X)]$ for 421 parameter combinations. The remaining 11 combinations for which the exclusion of nonstationary signals did not increase the contrast between the focal and nonfocal signals found by the randomness test are listed in Table I. We see that these exceptions are mostly found at small values of H in combination with small values of m and τ, or at high values of H. Accordingly many of the exceptions listed in Table I can be attributed to an insufficient reconstruction of the dynamics caused by too short an embedding window.

TABLE I. Exceptional parameter combinations for which we obtain the atypical result $D[p_f(\mathcal{U}_1^X|\mathcal{S}_0), p_n(\mathcal{U}_1^X\mathcal{S}_0)] < D[p_f(\mathcal{U}_1^X), p_n(\mathcal{U}_1^X)]$.

H	m	τ	k
1	2	2	1
1	2	4	1
1	2	4	5
2	2	1	1
2	2	2	1
2	2	1	5
2	2	2	5
2	2	1	10
32	6	2	1
32	2	8	1
32	8	8	10

As third example we show the dependence of $p(\mathcal{B}_1|\mathcal{U}_0^X\mathcal{U}_0^Y), p(\mathcal{B}_1|\mathcal{U}_1^X\mathcal{U}_0^Y)$ and $p(\mathcal{B}_1|\mathcal{U}_1^X\mathcal{U}_1^Y)$ on m, τ, and k, again for a fixed $H = 4$ sampling times [Fig. 7(c)]. Without exception we find $p(\mathcal{B}_1|\mathcal{U}_0^X\mathcal{U}_0^Y) < p(\mathcal{B}_1|\mathcal{U}_1^X\mathcal{U}_0^Y)$, and $p(\mathcal{B}_1|\mathcal{U}_0^X\mathcal{U}_0^Y) < p(\mathcal{B}_1|\mathcal{U}_1^X\mathcal{U}_1^Y)$ is found with the only exception of $H = 4, m = 2, \tau = 4, k = 5$. Again this exception is found at small values of the embedding window. Accordingly, the fact that the rejection of the randomness and nonlinear-independence test are correlated is found across almost all parameters shown in Fig. 7(c). The additional finding that the outcome of both randomness tests being rejected increases the rejection probability of the nonlinear-independence test more than the outcome of only one randomness test being rejected ($p(\mathcal{B}_1|\mathcal{U}_1^X\mathcal{U}_0^Y) < p(\mathcal{B}_1|\mathcal{U}_1^X\mathcal{U}_1^Y)$) is found only for high enough values of the embedding window $(m-1)\tau$. This again points to the importance of an embedding window of sufficient length to properly reconstruct complex dynamics. Very similar results are found for prediction horizons higher or lower than $H = 4$.

APPENDIX B: CONFIDENCE INTERVALS

We use

$$\left[p - 1.96\sqrt{\frac{p(1-p)}{n}},\ p + 1.96\sqrt{\frac{p(1-p)}{n}}\right] \quad \text{(B1)}$$

as approximation to the 95% confidence interval for our probability estimates $p(\cdot)$. Suppose that we carry out n Bernoulli trials with parameter p_0. That means we make n independent trials where the probability to get a positive result in an individual trial is p_0 and to get a negative result is $1 - p_0$. Then the total count of positive trial results follows

a binomial distribution with mean np_0 and standard deviation $\sqrt{np_0(1-p_0)}$. For large np_0 this binomial distribution can be approximated with a normal distribution with the same mean and standard deviation. Accordingly, the fraction of positive trial results, corresponding to our p, follows approximately a normal distribution with mean p_0 and standard deviation $\sigma_0 = \sqrt{\frac{p_0(1-p_0)}{n}}$. Replacing the unknown p_0 by our p derived from the sample to estimate σ_0 and using the inverse of the standard normal distribution function $\Phi^{-1}(1-0.05/2) = 1.96$ results in the confidence interval boundaries given above.

Confidence intervals for the quantity $D(p_1, p_2)$ in Eq. (6) can be derived as follows. Let u and v be two independent normally distributed random variables with means μ_u, μ_v and standard deviations σ_u, σ_v. Let the random variable d be defined as

$$d = \frac{u-v}{u+v}, \tag{B2}$$

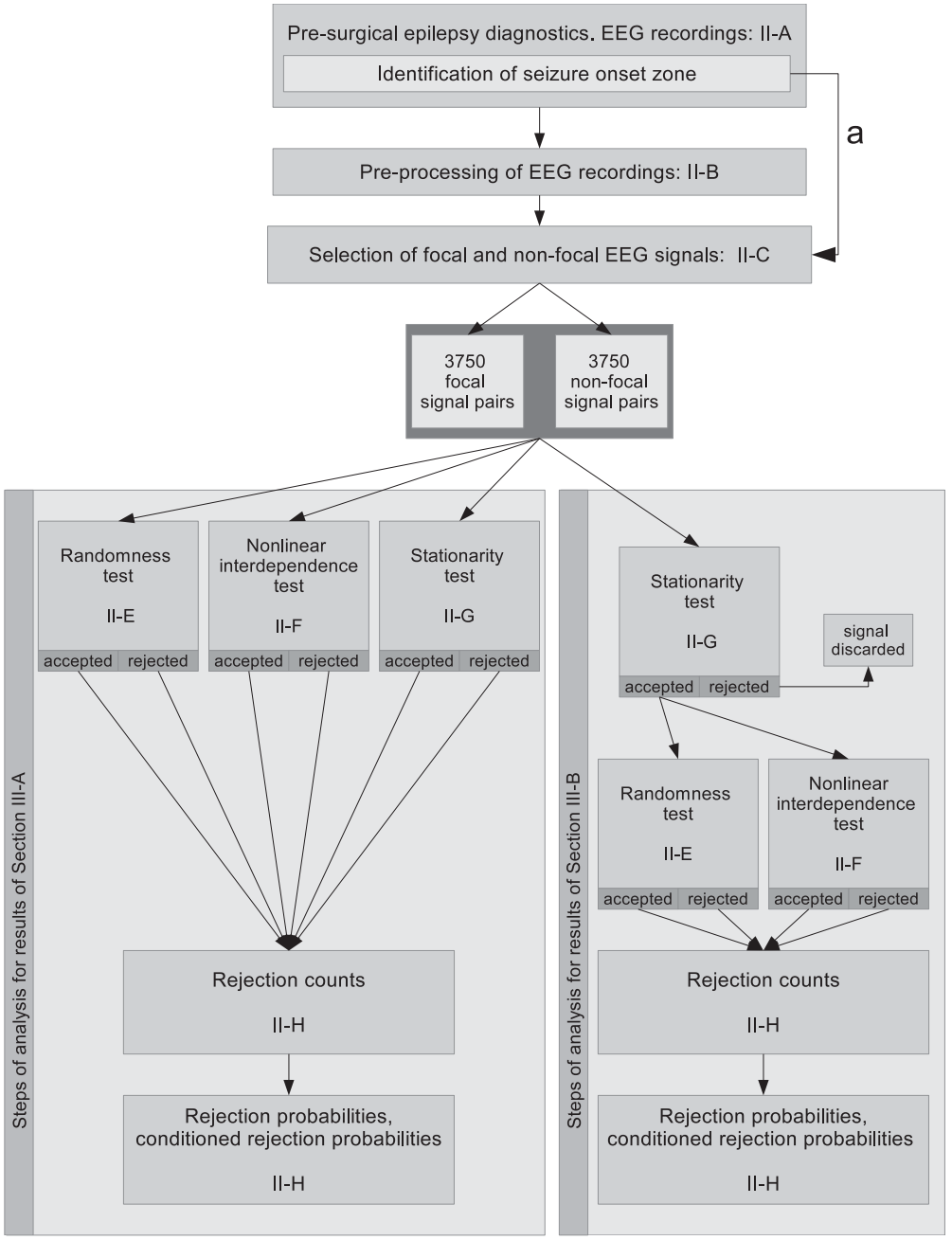

FIG. 8. Flowchart of the steps of analysis. The numbers refer to sections of the main manuscript. The unidirectional arrows emphasize that our analysis is sequential. Lower blocks have no influence on upper blocks. In particular, our data analysis has no influence on any aspect of the EEG recordings carried out during the presurgical epilepsy diagnostics. Likewise our results have no influence on the selection of the EEG signals. Arrow (a) is used to emphasize that the definition of the focal and nonfocal channels is fully determined through the information derived prior to and independent from our data analysis during the presurgical epilepsy diagnostics.

the distribution function $F(z)$ of which is

$$F(z) = P\left(\frac{u-v}{u+v} \leqslant z\right) \tag{B3}$$

$$= P[u - v \leqslant z(u+v)] \tag{B4}$$

$$= P[u(1-z) - v(1+z) \leqslant 0]. \tag{B5}$$

The left-hand side of the inequality is again a normally distributed random variable, which we denote by w. Its mean and standard deviation are, respectively,

$$\mu_w = \mu_u(1-z) + \mu_v(1+z), \tag{B6}$$

$$\sigma_w = \sqrt{\sigma_u^2(1-z)^2 + \sigma_v^2(1+z)^2}. \tag{B7}$$

Hence the random variable $W = \frac{w-\mu_w}{\sigma_w}$ follows a standard normal distribution with zero mean and unit variance, and

$$F(z) = P\big(w \leqslant 0\big) \tag{B8}$$

$$= P\left(W \leqslant -\frac{\mu_w}{\sigma_w}\right) \tag{B9}$$

$$= \Phi\left(-\frac{\mu_w}{\sigma_w}\right). \tag{B10}$$

Setting $F(z) = 0.975$ and $F(z) = 0.025$ we get

$$\Phi^{-1}(0.025) = -1.96 = -\frac{\mu_w}{\sigma_w}, \tag{B11}$$

$$\Phi^{-1}(0.975) = 1.96 = -\frac{\mu_w}{\sigma_w}. \tag{B12}$$

Solving Eq. (B11) for z, using (B6)–(B7) and defining $C = 1.96^2$, yields

$$\frac{\mu_u^2 - \mu_v^2 - C\big(\sigma_u^2 - \sigma_v^2\big) \pm \sqrt{4C\big(\mu_u^2\sigma_v^2 + \mu_v^2\sigma_u^2 - C\sigma_u^2\sigma_v^2\big)}}{(\mu_u + \mu_v)^2 - C\big(\sigma_u^2 + \sigma_v^2\big)}. \tag{B13}$$

Due to the symmetry of the problem, solving Eq. (B12) for z leads to the same pair of solutions. The resulting z_1 and z_2 are, respectively, the upper and lower boundary of the confidence interval for d.

APPENDIX C: STEPS OF ANALYSIS: FLOW CHART

A flowchart of the different steps of analysis is shown in Fig. 8.

[1] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. (Cambridge University Press, Cambridge, 2004).
[2] C. J. Stam, Clin. Neurophysiol. **116**, 2266 (2005).
[3] E. Pereda, R. Quian Quiroga, and J. Bhattacharya, Prog. Neurobiol. **77**, 1 (2005).
[4] K. Lehnertz, J. Biol. Phys. **34**, 253 (2008).
[5] M. Casdagli, L. Iasemides, R. Savit, R. Gilmore, S. Roper, and J. Sackellares, Electroenceph. Clin. Neurophysiol. **102**, 98 (1997).
[6] R. G. Andrzejak, G. Widman, K. Lehnertz, P. David, and C. E. Elger, Epilepsy Res. **44**, 129 (2001).
[7] R. G. Andrzejak, F. Mormann, G. Widman, T. Kreuz, C. E. Elger, and K. Lehnertz, Epilepsy Res. **69**, 30 (2006).
[8] R. G. Andrzejak, D. Chicharro, K. Lehnertz, and F. Mormann, Phys. Rev. E **83**, 046203 (2011).
[9] G. Varotto, L. Tassi, S. Franceschetti, R. Spreafico, and F. Panzica, Neuroimage **61**, 591 (2012).
[10] D. Kaplan and R. Cohen, Circ. Res. **67**, 886 (1990).
[11] J. Pijn, J. van Neerven, A. Noest, and F. Lopes da Silva, Clin. Neurophysiol. **79**, 371 (1991).
[12] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, Physica D **58**, 77 (1992).
[13] T. Schreiber and A. Schmitz, Phys. Rev. Lett. **77**, 635 (1996).
[14] X. Luo, T. Nakamura, and M. Small, Phys. Rev. E **71**, 026230 (2005).
[15] D. Prichard and J. Theiler, Phys. Rev. Lett. **73**, 951 (1994).
[16] T. Schreiber and A. Schmitz, Physica D **142**, 346 (2000).
[17] K. T. Dolan and A. Neiman, Phys. Rev. E **65**, 026108 (2002).
[18] R. G. Andrzejak, A. Kraskov, H. Stögbauer, F. Mormann, and T. Kreuz, Phys. Rev. E **68**, 066202 (2003).
[19] M. Thiel, M. Romano, J. Kurths, M. Rolfs, and R. Kliegl, Europhys. Lett. **75**, 535 (2006).
[20] X. Li, D. Cui, P. Jiruska, J. Fox, X. Yao, and J. Jefferys, J. Neurophysiol. **98**, 3341 (2007).
[21] M. Müller, G. Baier, C. Rummel, and K. Schindler, Europhys. Lett. **84**, 10009 (2008).
[22] M. A. Kramer, U. T. Eden, S. S. Cash, and E. D. Kolaczyk, Phys. Rev. E **79**, 061916 (2009).
[23] C. Rummel, M. Müller, G. Baier, F. Amor, and K. Schindler, J. Neurosci. Meth. **191**, 94 (2010).
[24] C. Wilke, W. van Drongelen, M. Kohrman, and B. He, Epilepsia **51**, 564 (2010).
[25] S. Bialonski, M. Wendler, and K. Lehnertz, PloS ONE **6**, e22826 (2011).
[26] T. Schreiber, Phys. Rev. Lett. **80**, 2105 (1998).
[27] C. Rummel, E. Abela, M. Müller, M. Hauf, O. Scheidegger, R. Wiest, and K. Schindler, Phys. Rev. E **83**, 066215 (2011).
[28] N. Burioka, M. Miyata, G. Cornelissen, F. Halberg, T. Takeshima, D. T. Kaplan, H. Suyama, M. Endo, Y. Maegaki, T. Nomura, Y. Tomita, K. Nakashima, and E. Shimizu, Clin. EEG Neurosci. **36**, 188 (2005).
[29] A. Aarabi, F. Wallois, and R. Grebe, Physica D **1188**, 207 (2008).
[30] H. Jing and M. Takigawa, Biol. Cybern. **83**, 391 (2000).
[31] G. Ouyang, X. Li, C. Dang, and D. A. Richards, Phys. Rev. E **79**, 041146 (2009).
[32] S. Bialonski, M. Wendler, and K. Lehnertz, Epilepsy Res. **99**, 78 (2012).
[33] H. G. Wieser, W. T. Blume, G. Fish, E. Goldensohn, A. Hufnagel, D. King, M. R. Sperling, and H. Lüders, Epilepsia **42**, 282 (2001).
[34] C. Rummel, G. Baier, and M. Müller, J. Neurosci. Meth. **166**, 138 (2007).

[35] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, Phys. Rev. E **64**, 061907 (2001).
[36] D. Chicharro and R. G. Andrzejak, Phys. Rev. E **80**, 026217 (2009).
[37] J. Arnhold, K. Lehnertz, P. Grassberger, and C. E. Elger, Physica D **134**, 419 (1999).
[38] A. Cenys, G. Lasiene, and K. Pyragas, Physica D **52**, 332 (1991).
[39] S. J. Schiff, P. So, T. Chang, R. E. Burke, and T. Sauer, Phys. Rev. E **54**, 6708 (1996).
[40] R. Quian Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger, Phys. Rev. E **65**, 041903 (2002).
[41] R. Quian Quiroga, J. Arnhold, and P. Grassberger, Phys. Rev. E **61**, 5142 (2000).
[42] M. C. Romano, M. Thiel, J. Kurths, and C. Grebogi, Phys. Rev. E **76**, 036211 (2007).
[43] J. Wang and H. Wieser, Epilepsia **35**, 495 (1994).
[44] T. Dikanev, D. Smirnov, R. Wennberg, J. Perez Velazquez, and B. Bezruchko, Clin. Neurophysiol. **116**, 1796 (2005).
[45] K. Schindler, H. Gast, L. Stieglitz, A. Stiebel, M. Hauf, R. Wiest, L. Mariani, and C. Rummel, Epilepsia **52**, 1771 (2011).
[46] R. G. Andrzejak and T. Kreuz, Europhys. Lett. **96**, 50012 (2011).
[47] All material is linked from the first author's home page: http://www.dtic.upf.edu/∼ralph/. Detailed description of the organization and formats of the files are provided there.