

## Detecting determinism from point processes

Ralph G. Andrzejak,<sup>1</sup> Florian Mormann,<sup>2</sup> and Thomas Kreuz<sup>3</sup>

<sup>1</sup>*Universitat Pompeu Fabra, Department of Information and Communication Technologies, E-08018 Barcelona, Spain*

<sup>2</sup>*Department of Epileptology, University of Bonn, D-53105 Bonn, Germany*

<sup>3</sup>*Institute for Complex Systems, CNR, I-50119 Sesto Fiorentino, Italy*

(Received 9 June 2014; published 2 December 2014)

The detection of a nonrandom structure from experimental data can be crucial for the classification, understanding, and interpretation of the generating process. We here introduce a rank-based nonlinear predictability score to detect determinism from point process data. Thanks to its modular nature, this approach can be adapted to whatever signature in the data one considers indicative of deterministic structure. After validating our approach using point process signals from deterministic and stochastic model dynamics, we show an application to neuronal spike trains recorded in the brain of an epilepsy patient. While we illustrate our approach in the context of temporal point processes, it can be readily applied to spatial point processes as well.

DOI: [10.1103/PhysRevE.90.062906](https://doi.org/10.1103/PhysRevE.90.062906)

PACS number(s): 05.45.Tp, 84.35.+i, 87.19.xm, 87.19.L–

Point process signals are ubiquitous in nature. The underlying dynamics is often intrinsically stochastic. For example, radioactive decay is adequately described by a stochastic Poisson process. Other point process signals can be assumed to contain stochastic and deterministic components. A prominent example is the sequences of action potentials, generally referred to as spike trains, used by the nervous system to transmit information. Here the detection of deterministic structure can be key for a thorough characterization of neuronal dynamics. Motivated by this example from neuroscience, but without loss of generality, we use the term spike train to refer to point process signals. We here introduce an approach to detect determinism from spike trains. It has the important advantage over existing approaches [1,2] that it can be targeted at whatever feature one considers relevant for the detection of deterministic structure, such as the spike rate or spike timing.

For noise-free deterministic dynamics a given present state unambiguously determines the dynamics' future devolution. For smooth deterministic dynamics, similar instantaneous states remain similar in their immediate future. A number of classical tests for determinism (e.g., [3–7]) are therefore based on quantifying the dynamics' degree of predictability. For this purpose, a prediction of the future evolution of a reference state is derived from the evolution of states which are similar to this reference state. In these approaches, similarity is assessed by spatial proximity in reconstructed state spaces. Importantly, it is well established that predictability is a necessary but not sufficient criterion for determinism. This is because also for stochastic processes with memory, similar instantaneous states remain similar in their immediate future. Spurious detections of predictability might furthermore occur for stochastic signals that have no memory but are, for instance, nonstationary. It is therefore essential to use surrogate signals to test whether the results of some measure for predictability are consistent with the null hypothesis of a stochastic process [8]. Only in combination with surrogate null hypothesis testing can tests for predictability be used as a test for determinism. A complementary approach to detect and characterize deterministic structure is to inspect quantities such as finite size Lyapunov exponents or epsilon entropies estimated from the signals (e.g., [9]).

Given a spike train, one can use inter-spike-interval embeddings to reconstruct the generating dynamics [1,2]. Here present states are formed by vectors whose components contain the intervals between the most recent spikes. Each vector is taken to represent the state of the dynamics at the time of the spike closing the last interval. The sequence of these vectors across all spikes in the spike train yields the inter-spike-interval embedding. In analogy to delay coordinate embeddings applied to time-continuous signals [10], inter-spike-interval embeddings are equivalent to the state space of the time-continuous dynamics generating the spike train [1,2,11]. To illustrate this point, Sauer [1,2] used Euclidean inter-state-distance matrices calculated from inter-spike-interval embeddings as input for a nonlinear prediction error [5]. This approach allowed Sauer to distinguish between spike trains of deterministic versus stochastic origin.

We here introduce a rank-based nonlinear predictability score for spike trains, termed  $\mathcal{S}$ , and apply it to original and surrogate spike trains. The important difference to the classic approach above is that we bypass the state space reconstruction. We assess the similarity between instantaneous states, as represented by short segments of the spike train, directly by means of spike train distances [12]. Importantly, here one can use any measure of spike train distance. For example, one can apply metrics that focus either on spike rates or on spike timing. We here show that this adaptability is not only an important conceptual advantage, but is indeed crucial for achieving optimal sensitivity for determinism across different dynamics.

To calculate  $\mathcal{S}$  [13], the spike times  $\{t_l\}_{l=1,\dots,L}$ , are scaled to the dimensionless interval  $[t_1 = 0, t_L = 1]$ . The spike train is then segmented into overlapping intervals of length  $q$ . Each segment represents the state of the dynamics at the end of this segment. The step size between subsequent segments is  $s$ , with  $s \leq q \ll 1$ . Accordingly, the  $i$ -th segments span  $[(i-1)s, (i-1)s + q]$  for  $i = 1, \dots, N$ , where  $N$  is obtained by rounding down  $\frac{1-q}{s} + 1$ . To quantify the dissimilarity  $d_{ij}$  between the pair of states represented by the segments  $i$  and  $j$ , we apply spike train distances to this pair of segments. Across all pairs this yields the inter-state-distance matrix  $D_{ij}$ . As spike train distances, we use the ISI distance [14,15] and SPIKE distance

[15–17]. Given two spike trains, the ISI distance quantifies the average instantaneous similarity between the inter-spike-interval profiles while the precise timing of spikes is irrelevant. In contrast, the SPIKE distance quantifies the degree to which the spikes occur at similar times across the two spike trains.

Given the matrix  $D_{ij}$ , one can readily quantify the dynamic's predictability for the horizon  $h$ . For each reference state with index  $i_0 = 1, \dots, N - h$ , we denote the indices of its  $k$  nearest neighbors by  $\{j_{0,r}\}_{(r=1,\dots,k)}$ . Here we impose  $j_{0,r} \leq N - h$  and  $|i_0 - j_{0,r}| > w$ , with decorrelation window  $w$  [18]. Accordingly,  $\{j_{0,r}\}_{(r=1,\dots,k)}$  are the indices of the  $k$  smallest entries in row  $i_0$  of  $D_{ij}$ , ignoring the entries in the main diagonal, in the  $w$  first off-diagonal pairs, and in the last  $h$  columns. Furthermore, we denote by  $g_{i_0}^j$  the rank which  $d_{i_0,j}$  assumes in a sorted list of all distances included in the set  $\{d_{i_0,n}\}_{n=1,\dots,N;|i_0-n|>w}$ . This set corresponds to row  $i_0$  of  $D_{ij}$ , now only ignoring the entries in the main diagonal and the first  $w$  off-diagonal pairs. Accordingly, the number of distances in this set is  $M_{i_0} = N - 2w - 1$  for  $w < i_0 < N - w + 1$ . Below and above this range,  $M_{i_0}$  increases linearly and reaches  $M_{i_0} = N - w - 1$  at  $i_0 = 1$  and  $i_0 = N$ .

As stated above, similar instantaneous states of deterministic dynamics remain similar in their immediate future. To quantify the resulting predictability, we at first determine the mean rank of the  $k$  nearest neighbors  $j_{0,r}$  of the reference point  $i_0$  after  $h$  time steps have passed:  $R_{i_0} = \frac{1}{k} \sum_{r=1}^k g_{i_0+h}^{j_{0,r}+h}$ . For complete predictability,  $\{j_{0,r} + h\}_{(r=1,\dots,k)}$  are still the indices of the nearest neighbors of  $i_0 + h$ , and we get  $R_{i_0} = \frac{k+1}{2}$ . We denote this lower boundary, which is independent of  $i_0$ , by  $R^L = \frac{k+1}{2}$ . In contrast, for no predictability, our nearest neighbors will have lost all proximity  $h$  time steps later. Accordingly,  $g_{i_0+h}^{j_{0,r}+h}$  are just  $k$  random samples from a uniform distribution on the integers  $1, \dots, M_{i_0}$ . In consequence, for no predictability,  $R_{i_0}$  has an expected value of  $R_{i_0}^U = \frac{M_{i_0}+1}{2}$  [19]. Therefore, using a general normalization principle [20,21], we define the rank-based predictability score:

$$\mathcal{S} = \frac{1}{N - h} \sum_{i_0=1}^{N-h} \frac{R_{i_0}^U - R_{i_0}}{R_{i_0}^U - R^L}. \quad (1)$$

For complete predictability we get  $R_{i_0} = R^L$ , and  $\mathcal{S}$  attains its maximal value of 1. In contrast, for no predictability  $\mathcal{S}$  has an expected value of zero since here  $R_{i_0}$  has an expected value of  $R_{i_0}^U$ . Furthermore, as a consequence of the central limit theorem, for no predictability  $\mathcal{S}$  is normally distributed since here it reduces to a sum of samples of an identically distributed random variable. We use  $\mathcal{S}_{\text{ISI}}$  and  $\mathcal{S}_{\text{SPIKE}}$  whenever we distinguish which spike train distance was used.

We use five models of spike trains derived from simulated systems. Four models, denoted by A–D, are generated by deterministic dynamics to assess the sensitivity of our approach and its robustness against noise. To assess its specificity, we test whether we falsely detect deterministic structure in stochastic spike trains. Therefore, for model E we generate spike trains by a stochastic dynamics. Furthermore, we analyze real neuronal spike trains recorded from a patient with epilepsy. To stress the applicability of our approach to experimental data, we use spike trains of only 500 spikes for all models and the neuronal data.

For model A [2,12] an integrate-and-fire process  $S(t)$  is driven by the  $x(t)$  variable of a Lorenz dynamics [ $\dot{x}(t) = 10(y(t) - x(t))$ ;  $\dot{y}(t) = 28x(t) - y(t) - x(t)z(t)$ ;  $\dot{z}(t) = x(t)y(t) - \frac{8}{3}z(t)$ ]. We define the initial time as  $t_1 = 0$  and set  $S(t_1) = 0$ . Then for  $l = 2, \dots, 500$ , we integrate  $S(t) = \int_{t'=\tau_{l-1}}^t (x(t') + 25)dt'$ . The time when  $S(t)$  crosses the threshold of  $\Theta = 12$  is used as spike time  $t_l$ ,  $S(t_l)$  is reset to zero, and the integration is restarted. For model B [12,22] we define spike times directly from upward threshold crossings ( $\Theta = 27$ ) of the variable  $z(t)$  of the Lorenz dynamics. For models C and D we use a Hindmarsh-Rose neuron ( $\dot{x}(t) = y(t) + 3x^2(t) - x^3(t) - z(t) + J$ ;  $\dot{y}(t) = 1 - 5x^2(t) - y(t)$ ;  $\dot{z}(t) = 0.0021[-z(t) + 4(x(t) + 1.6)]$ ). In contrast to the Lorenz dynamics, this dynamics exhibits intrinsic spiking. The  $x(t)$  variable shows prominent spikes which we can readily detect by the upward crossing of the threshold  $\theta = 0.6$  [12,23]. The neuron's dynamical regime depends on the variable  $J$ . For model C we set  $J = 3.30$ , which leads to aperiodic spiking. For model D we use  $J = 3.28$ , for which the spikes are grouped in aperiodic bursts. Model E is derived in the same way as model A. However, the deterministic Lorenz variable is replaced by a stochastic signal to drive the integrate-and-fire process [1]. This stochastic signal is constrained to have the same power spectrum and amplitude distribution as  $x(t)$  but is otherwise random [24]. Accordingly, the resulting spike trains are stochastic.

To test the robustness of our approach, we superimpose the model spike trains with noise. We delete a percentage  $\gamma$  of randomly selected spikes and afterwards insert the same number of spikes uniformly randomly distributed across the duration of the spike train. This reflects that for experimental data, the detection algorithm can have false positive and false negative spike detections, or it can assign wrong times to true positive spike detections.

We furthermore analyze a recording of neuronal spiking from the medial temporal lobe of an epilepsy patient. This recording was performed prior to and independently from this study at the University of Bonn, Germany, while the patient underwent monitoring prior to epilepsy surgery (see [17] for details). During the 10-min recording, the patient was awake and at rest, no seizure occurred. The recording contained spike trains that were classified in pre-analysis to originate each from a single neuron as well as spike trains that reflect the mixed activity of multiple neurons (multi-unit activity). We analyzed only those 18 single neuron spike trains that contained at least 500 spikes. Of these spike trains we took the first 500 spikes.

To generate spike train surrogates, we shuffle the order of the inter-spike-intervals. This corresponds to the null hypothesis that the sequence of inter-spike-intervals is stationary and uncorrelated. Since the original inter-spike-interval distribution is maintained, no assumption is made about this distribution. In particular, it is neither assumed nor excluded that the spike train is generated by a Poisson process which would result in an exponential inter-spike-interval distribution.

Apart from fixing the number of spikes to  $L = 500$ , we use fixed parameters across all model systems and the neuronal data. We define the length of the spike train to 1 dimensionless time unit, and use  $q = 0.01$  and  $s = 0.001$  for the segmentation. The smaller  $s$ , the better the sampling by the inter-state-distance matrix. Very small values, however,

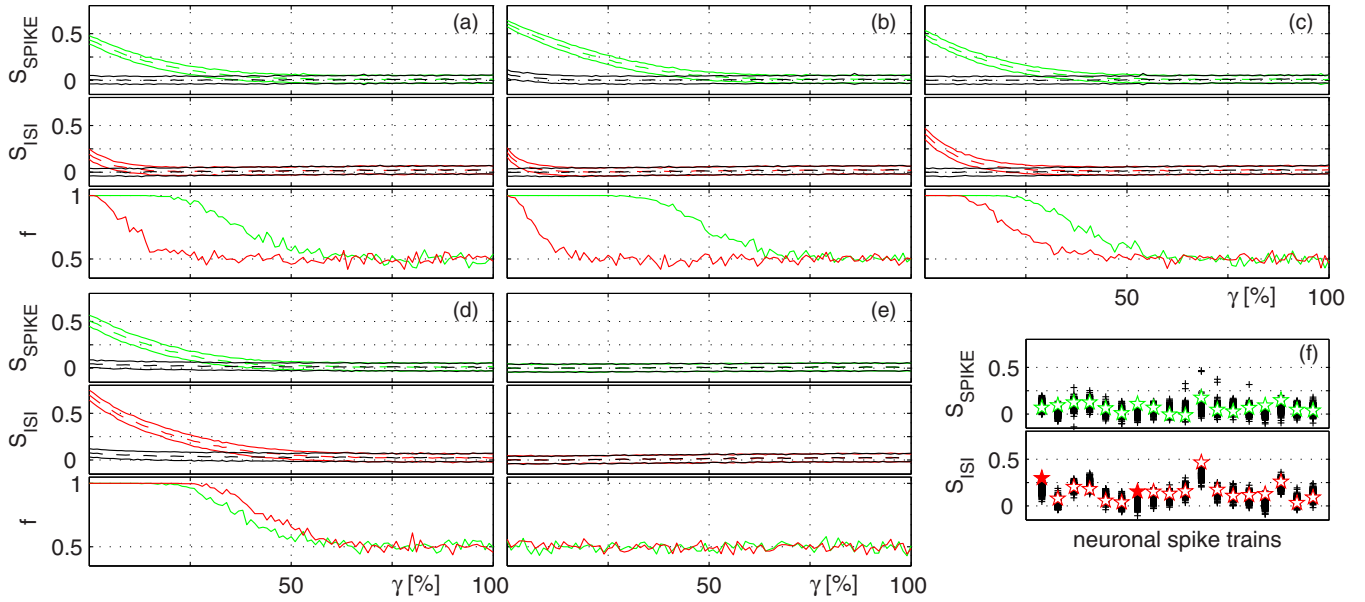


FIG. 1. (Color online) Results for  $S_{\text{SPIKE}}$  and  $S_{\text{ISI}}$  for the original spike trains are shown in green (light grey) and red (dark grey), respectively. Results for the corresponding surrogates are shown in black [lower bands in (a)–(e), crosses in (f)]. (a)–(e) dependence of results on the noise percentage  $\gamma$  for models A–E, respectively. For each model and every  $\gamma$  value we generated 300 independent spike trains realizations. From each realization we generated one surrogate spike train. Bands in the two upper parts show averages  $\pm$  one standard deviation of  $S$  across the 300 original spike trains and across the 300 surrogates. Solid lines in the lower parts show the fraction  $f$  of realizations for which the  $S$  for the original spike train was higher than the one for its surrogate. (f) results for the neuronal spiking data (pentagons) and 199 surrogates (crosses). Filled pentagons indicate cases for which the  $S$  values of the original exceeds all surrogates.

increase the computational load. The parameter  $q$  determines the average number of spikes per segment. At  $q = 0.01$  we have an average of five spikes per segment. The prediction horizon is an integer multiple of  $s$  and should be at least  $q$  to avoid overlap between predicting and predicted segments. We use  $h = 0.014 > q$  to reduce the influence of linear correlations between subsequent segments. The decorrelation window  $w$  serves to avoid that temporally close segments are selected as nearest neighbors [18]. We use a high but otherwise arbitrary value of  $w = 0.05$ . Finally, we fix  $k = 1$  nearest neighbor. Exploring the parameter space, we found these values to work well across all model systems and the neuronal data. We deliberately refrained from any extensive parameter optimization for individual settings. Neither were the data selected as to optimize our results. The deterministic models A–D were taken over from [12] where they were studied in the context of detecting directional couplings. The stochastic model E was created for the present study, since in [12] the specificity was tested using independent but deterministic spike trains. The particular patient and recording of neuronal spike trains was selected at random from a larger pool of data.

We first consider results obtained for spike trains derived from models A–E in dependence on the noise level [Fig. 1(a)–1(e)]. The  $S_{\text{SPIKE}}$  values for the original spike trains exceed the surrogate results up to substantial noise levels for all deterministic models (A–D). Even if 25%–50% of the spikes are relocated to random times,  $S_{\text{SPIKE}}$  correctly detects the deterministic structure. A more heterogeneous picture is obtained for  $S_{\text{ISI}}$ . For noise-free dynamics, also  $S_{\text{ISI}}$  correctly detects the deterministic structure. However, for the models A–C already at low noise levels,  $S_{\text{ISI}}$  loses sensitivity. In

contrast, for spike trains from the Hindmarsh-Rose neuron in an irregular bursting dynamics (model D),  $S_{\text{ISI}}$  performs better and even slightly exceeds the already high robustness of  $S_{\text{SPIKE}}$ . For the spike trains generated by a stochastic dynamics (model E), the original and the surrogates match across all noise levels for both  $S_{\text{SPIKE}}$  and  $S_{\text{ISI}}$ . Hence, these results provide evidence for the sensitivity and specificity of our approach for deterministic structure in spike trains.

The mean value of  $S$  for stochastic signals, as represented by the highest noise levels or by the surrogates, sometimes exceeds zero, although the expected value of  $S$  for no predictability is zero. This is due to linear correlations between neighboring elements in the inter-state-distance matrices. These correlations arise since the distances used here are derived from piecewise constant ( $S_{\text{ISI}}$ ) or piecewise linear ( $S_{\text{SPIKE}}$ ) profiles. Accordingly, unless there is a spike exactly at the border between two segments, the profile values match at these borders. This effect has no impact since it affects the original and the surrogate results in the same way. If needed, it can be suppressed using higher  $q$  values.

For each of the 18 neuronal spike trains we generated 199 surrogates. For  $S_{\text{SPIKE}}$  the value for the original spike train never exceeds the range of the surrogates. In contrast, in two cases the  $S_{\text{ISI}}$  obtained for the original spike train is higher than for all surrogates [Fig. 1(f)]. If the null hypothesis was correct for all spike trains, there would be a probability of only 0.0036 to get two or more rejections. The overall low rejection percentage should not be misinterpreted as corresponding to the sensitivity of our test to detect deterministic structure from real neuronal data. Rather it can indicate that different types of neurons sampled in the recordings in fact have

distinct dynamical properties. For example, excitatory versus inhibitory neurons (e.g., [25]) may show different degrees of non-randomness.

Apart from this physiological aspect, the disease epilepsy can have an important impact on the characteristics of neuronal dynamics. Earlier studies of intracranial electroencephalographic (EEG) recordings from epilepsy patients during the seizure-free interval showed that while recordings from healthy brain areas were mostly consistent with the null model of a linear stochastic correlated process, recordings from brain areas generating epileptic seizures exhibit a more nonrandom behavior [26–28]. These findings suggest that neurons which are involved in the initiation of epileptic seizures may reveal an increased degree of nonrandomness also during the seizure-free period. Our approach could allow one to detect these putative ictogenic neurons, thereby contributing information useful for diagnostic purposes. It might also help to further discriminate between the recently described distinct electrophysiological phenotypes of neurons in hypothalamic hamartoma [29].

Importantly, the prominent differences between  $\mathcal{S}_{\text{ISI}}$  and  $\mathcal{S}_{\text{SPIKE}}$  show that deterministic structure manifests itself in different features of the spike trains across the different models and the neuronal data. In consequence, depending on the dynamics, the resulting predictability is better captured if the inter-state-distance matrices are calculated using one or the other spike train distance. Overall,  $\mathcal{S}_{\text{SPIKE}}$  performs better than  $\mathcal{S}_{\text{ISI}}$  in detecting nonrandomness in the model systems. On the other hand, for the neuronal data only  $\mathcal{S}_{\text{ISI}}$  led to rejections of the surrogate null hypothesis. It is the adaptability of our approach to these different settings that makes it very versatile and powerful. One can *a priori* choose a distance that focuses on what one considers to be the relevant feature of the spike train. Conversely, by comparing the outcome for

different distances, one can detect this distinctive feature *a posteriori*. Hence, this analysis can for example address the neural coding problem of distinguishing between a rate code and a temporal code. To scan for deterministic structure on specific time scales, one can use the Victor-Purpura spike train distance [30] which offers a corresponding continuous parameter (cf. [31]).

Further flexibility is offered by the choice of surrogates. For the sake of generality, we used very simple surrogates which test the null hypothesis that the sequence of inter-spike-intervals is stationary and uncorrelated. One can argue that this does not represent a plausible null model for neuronal spiking, since already an inhomogeneous Poisson process might reject this null hypothesis. However, it was not our aim to find an adequate null model for the investigated spike trains. Rather we used high values of  $h$  and  $w$  to minimize the influence of correlations between inter-spike-intervals. In future studies more elaborate surrogate types can be used that test for more general null models. These null models can be tailored to the particular type of experimental data and working hypothesis. Furthermore, apart from testing for predictability, any measure that is based on inter-state-distances can be used. One example is entropy estimates based on  $k$  nearest neighbor distances and information-theoretic measures derived from them [32].

R.G.A. acknowledges Grant No. FIS-2010-18204 of the Spanish Ministry of Education and Science. R.G.A. and F.M. acknowledge funding from the Volkswagen Foundation. T.K. acknowledges funding from the European Commission through the Marie Curie Initial Training Network ‘Neural Engineering Transformative Technologies (NETT)’, Project No. 289146. The authors thank N. Bozanic for optimizations of the spike train distances’ source codes.

- 
- [1] T. Sauer, *Phys. Rev. Lett.* **72**, 3811 (1994).  
 [2] T. Sauer, *Chaos* **5**, 127 (1995).  
 [3] J. D. Farmer and J. J. Sidorowich, *Phys. Rev. Lett.* **59**, 845 (1987).  
 [4] M. Casdagli, *Physica D* **35**, 335 (1989).  
 [5] G. Sugihara and R. May, *Nature* **344**, 734 (1990).  
 [6] M. Casdagli, *J. Roy. Stat. Soc., Ser B* **54**, 303 (1992).  
 [7] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, 1997).  
 [8] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, *Physica D* **58**, 77 (1992).  
 [9] M. Cencini, M. Falcioni, E. Olbrich, H. Kantz, and A. Vulpiani, *Phys. Rev. E* **62**, 427 (2000).  
 [10] F. Takens, in *Dynamical Systems and Turbulence*, edited by D. A. Rand and L.-S. Young (Springer-Verlag, Berlin, 1981), Vol. 898 of *Lecture Notes in Mathematics*, pp. 366–381.  
 [11] R. Hegger and H. Kantz, *EPL (Europhys. Lett.)* **38**, 267 (1997).  
 [12] R. G. Andrzejak and T. Kreuz, *EPL (Europhys. Lett.)* **96**, 50012 (2011).  
 [13] Source codes will be published at, <http://ntsa.upf.edu/downloads>.  
 [14] T. Kreuz, J. S. Haas, A. Morelli, H. D. I. Abarbanel, and A. Politi, *J. Neurosci. Meth.* **165**, 151 (2007).  
 [15] T. Kreuz, *Scholarpedia* **6**, 11934 (2011).  
 [16] T. Kreuz, D. Chicharro, M. Greschner, and R. G. Andrzejak, *J. Neurosci. Meth.* **195**, 92 (2011).  
 [17] T. Kreuz, D. Chicharro, C. Houghton, R. G. Andrzejak, and F. Mormann, *J. Neurophysiol.* **109**, 1457 (2013).  
 [18] J. Theiler, *Phys. Rev. A* **34**, 2427 (1986).  
 [19] Since  $k \ll M_{i_0}$ , we neglect that the  $k$  samples are taken without replacement.  
 [20] D. Chicharro and R. G. Andrzejak, *Phys. Rev. E* **80**, 026217 (2009).  
 [21] R. G. Andrzejak, A. Kraskov, H. Stögbauer, F. Mormann, and T. Kreuz, *Phys. Rev. E* **68**, 066202 (2003).  
 [22] M. Ding and W. Yang, *Phys. Rev. E* **55**, 2397 (1997).  
 [23] A. Kaiser and T. Schreiber, *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **14**, 1987 (2004).  
 [24] T. Schreiber and A. Schmitz, *Phys. Rev. Lett.* **77**, 635 (1996).  
 [25] M. J. Ison, F. Mormann, M. Cerf, C. Koch, I. Fried, and R. Quian Quiroga, *J. Neurophysiol.* **106**, 1713 (2011).

- [26] R. G. Andrzejak, G. Widman, K. Lehnertz, P. David, and C. E. Elger, *Epilepsy Res.* **44**, 129 (2001).
- [27] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, *Phys. Rev. E* **64**, 061907 (2001).
- [28] R. G. Andrzejak, K. Schindler, and C. Rummel, *Phys. Rev. E* **86**, 046206 (2012).
- [29] P. N. Steinmetz, S. D. Wait, G. P. Lekovic, H. L. ReKate, and J. F. Kerrigan, *Frontiers in Neurology* **4**, 210 (2013).
- [30] J. Victor and K. Purpura, *Network* **8**, 127 (1997).
- [31] D. Chicharro, T. Kreuz, and R. G. Andrzejak, *J. Neurosci. Meth.* **199**, 146 (2011).
- [32] A. Kraskov, H. Stögbauer, and P. Grassberger, *Phys. Rev. E* **69**, 066138 (2004).