Optimal estimation of recurrence structures from time series

PETER BEIM GRABEN[1], KRISTIN K. SELLERS[2,3], FLAVIO FRÖHLICH[2,3,4,5,6,7] and AXEL HUTT[8]

[1] *Bernstein Center for Computational Neuroscience - Berlin, Germany*
[2] *Department of Psychiatry, University of North Carolina at Chapel Hill - Chapel Hill, NC, USA*
[3] *Neurobiology Curriculum, University of North Carolina at Chapel Hill - Chapel Hill, NC, USA*
[4] *Department of Cell Biology and Physiology, University of North Carolina at Chapel Hill - Chapel Hill, NC, USA*
[5] *Department of Biomedical Engineering, University of North Carolina at Chapel Hill - Chapel Hill, NC, USA*
[6] *Neuroscience Center, University of North Carolina at Chapel Hill - Chapel Hill, NC, USA*
[7] *Department of Neurology, University of North Carolina at Chapel Hill - Chapel Hill, NC, USA*
[8] *German Meteorological Service - Offenbach am Main, Germany*

Abstract – Recurrent temporal dynamics is a phenomenon observed frequently in high-dimensional complex systems and its detection is a challenging task. Recurrence quantification analysis utilizing recurrence plots may extract such dynamics, however it still encounters an unsolved pertinent problem: the optimal selection of distance thresholds for estimating the recurrence structure of dynamical systems. The present work proposes a stochastic Markov model for the recurrent dynamics that allows for the analytical derivation of a criterion for the optimal distance threshold. The goodness of fit is assessed by a utility function which assumes a local maximum for that threshold reflecting the optimal estimate of the system's recurrence structure. We validate our approach by means of the nonlinear Lorenz system and its linearized stochastic surrogates. The final application to neurophysiological time series obtained from anesthetized animals illustrates the method and reveals novel dynamic features of the underlying system. We propose the number of optimal recurrence domains as a statistic for classifying an animals' state of consciousness.

Introduction. – Complex behavior is ubiquitous in nature, humanities and engineering. Often, complex dynamical systems are *recurrent* in the sense that certain regions of their available state space are frequently visited in the course of time [1,2]. This important property facilitates forecasting, modeling and control of dynamical systems. To visualize recurrent behavior, Eckmann *et al.* [3] suggested the recurrence plot (RP) technique that inspired the growing research field of recurrence quantification analyses (RQA). RP and RQA have found several applications in the physical sciences [4,5], medicine [6,7], social sciences [8,9] and engineering [10,11] (for further surveys, see [2,12]).

For a d-dimensional time series $\{\boldsymbol{x}_t\}, \boldsymbol{x}_t \in \mathbb{R}^d, t = 1, \ldots, N$, with N number of time steps, the RP is a graphical representation of the recurrence matrix

$$R_{ij} = \Theta(\varepsilon - ||\boldsymbol{x}_j - \boldsymbol{x}_i||). \tag{1}$$

The parameter ε is the distance threshold of two sampling points $\boldsymbol{x}_i, \boldsymbol{x}_j$, and Θ denotes the Heaviside step function. Proper selection of the threshold ε is of crucial importance to gain instructive RPs that reveal the underlying recurrence structure of the dynamics [2,12].

To solve the threshold selection problem, several heuristics have been suggested [2,12–14]. However, a proper optimization method should take the system's specific recurrence structure into account by explicitly modeling the alternating sequence of *recurrence domains* and *transients*. One viable way to do so, models the percolation transition toward the giant cluster in a recurrence network [15]. Another possibility is the detection of recurrence domains though *recurrence grammars* [13,14]. A recurrence grammar is comprised of rules $i \to j$ for the RP elements when $i > j$ and $R_{ij} = 1$, applied recursively to the sequence of time indices $s_t = t$. This map yields a symbolic dynamics s' which is a coarse-graining of the system's trajectory $\boldsymbol{x}_t$;

it transforms the recurrence structure into an alternating sequence of recurrence domains and transients. In this letter, we present a model for the system's recurrence structure in terms of a Markov chain obtained from its symbolic dynamics s'. Under the assumption that the optimal Markov chain describes transitions between several metastable states and a transient regime [14], we formulate three plausible criteria for the shape of the stochastic transition matrix leading to a novel utility function. Thus, its maximization over a range of distance thresholds ε yields an optimal estimate of the system's recurrence structure.

Method. – For easy comparison with our previously presented maximum entropy criterion [13,16], we consider the Lorenz system [17] as a well-known example for a system with a distinguished recurrence structure. Intuitively, the system exhibits two recurrence domains, namely the attractor's "wings" centered around its unstable foci and separated by several non-recurrent regimes, henceforth referred to as "transients". Therefore, in an optimal encoding the symbolic sequence s' obtained from a recurrence grammar would look like

$$s' = 00\ldots011\ldots100\ldots022\ldots200\ldots$$
$$111\ldots100\ldots022\ldots200\ldots \quad (2)$$

after applying the rewriting grammar [13,14,16]. *I.e.* we expect a number of 0's for some transient regime, followed by a relatively large number of 1's for the first wing, followed by a smaller number of 0's characterizing the transition to the second wing, which is hence reflected by a relatively large number of 2's, and so on. This symbolic dynamics is essentially a Markov chain with a rather simple transition graph depicted in fig. 1.

Figure 1 points out three important properties of the expected optimal encoding: i) The two recurrence domains (states 1 and 2) and the transient state 0 are modeled as metastable states with relatively large dwells, *i.e.* they exhibit mainly self-transitions. Hence the transition probabilities p_{00}, p_{11}, and p_{22} will be large compared to other transitions. ii) The transient state 0 acts as a "hub". Without further knowledge about the system's dynamics, it is reasonable to assume that the transition probabilities p_{10}, p_{20} from the hub and p_{01}, p_{02} into the hub are uniformly distributed according to a maximum entropy principle [13]. iii) There a no direct transitions between metastable recurrence domains 1 and 2. This model can be easily generalized to a higher number of recurrent states.

For a chosen threshold ε, the recurrence grammar yields a segmentation of the time series into $n(\varepsilon)$ symbols. Therefore, the appropriate model is an n-state Markov chain with number of recurrence domains (NRD) equal to $n - 1$. The transition matrix for an optimal recurrence grammar partition for $n > 1$ would be

$$\boldsymbol{P} = \begin{bmatrix} 1-(n-1)q & r & r & \cdots & r \\ q & 1-r & 0 & \cdots & 0 \\ q & 0 & 1-r & \cdots & 0 \\ \hdotsfor{5} \\ q & 0 & 0 & \cdots & 1-r \end{bmatrix}. \quad (3)$$
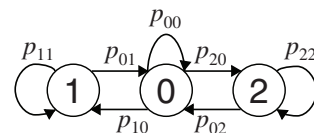

Fig. 1: Optimal Markov chain for recurrence grammar of the Lorenz attractor [17]. State 0 denotes the transient regime acting as a "hub". The metastable states 1 and 2, corresponding to the attractor's "wings", have considerably large dwells reflected by large self-transition probabilities p_{ii}.

For $n = 1$ the system exhibits only the transient state 0 and we set $\boldsymbol{P} = 1$. The matrix $\boldsymbol{P}$ reflects the three optimization criteria formulated above as follows: i) The trace $\operatorname{tr}\boldsymbol{P} = 1 + (n-1)(1-q-r)$ dominates all other transition probabilities. ii) The transition probabilities from the "hub" $q = p_{i0}$ and into the "hub" $r = p_{0j}$ for $i, j > 0$ are uniformly distributed. iii) As $\boldsymbol{P}$ is a stochastic matrix the column sums are $1-(n-1)q+(n-1)q = 1$ and $r+(1-r)=1$, the latter term excluding direct transitions between metastable states.

The three optimization criteria eventually lead to the desired utility function: i) Maximization of the recurrent self-transitions is achieved by maximizing the trace of the transition matrix $\operatorname{tr}\boldsymbol{P}$. ii) The maximum entropy principle is satisfied by renormalization of transition probabilities of the first row and of the first column of $\boldsymbol{P}$ after neglecting p_{00}, *i.e.*, $p'_{0j} = p_{0j}/\sum_{j=1}^{n-1} p_{0j}$ for the first row and $p'_{i0} = p_{i0}/\sum_{i=1}^{n-1} p_{i0}$ for the first column. Then,

$$h_r = -\frac{1}{\log(n-1)} \sum_{j=1}^{n-1} p'_{0j} \log p'_{0j},$$
$$h_c = -\frac{1}{\log(n-1)} \sum_{i=1}^{n-1} p'_{i0} \log p'_{i0}. \quad (4)$$

For the optimal Markov matrix, eq. (3), $h_r = h_c = 1$ (for $n = 1$ we trivially define $h_r = h_c = 0$, as there are no transitions). iii) Simultaneously maximizing the trace and the entropies of the first row and column of $\boldsymbol{P}$ also suppresses transitions between any two recurrence domains due to the normalization condition for stochastic matrices with $\sum_{i=0}^{n-1} p_{ij} = 1$. Finally, in the limit $q, r \to 0$, the optimal transition matrix in eq. (3) turns into the unit matrix $\boldsymbol{I}$ with $\operatorname{tr}\boldsymbol{I} = n$. Consequently, the utility function

$$u(\varepsilon) = \frac{1}{n+2}\Big[\operatorname{tr}\boldsymbol{P}(\varepsilon) + h_r(\varepsilon) + h_c(\varepsilon)\Big] \quad (5)$$

with $0 \le u(\varepsilon) \le 1$ allows us to define the optimization criterion

$$\varepsilon^* = \arg\max_\varepsilon u(\varepsilon). \quad (6)$$

The parameter ε^* is the optimal threshold for which the recurrence structure detected from the time series most resembles the Markov chain model. In addition, one may call $u(\varepsilon^*)$ the degree of resemblance of the time series' recurrence structure with the Markov chain model. The number

of recurrence domains (NRD), $n(\varepsilon^*) - 1$, characterizes the "complexity" of the metastable transition model.

For illustration, consider two interesting limiting cases. As long as the threshold ε remains smaller than the smallest distance between two sampling points, all points are isolated and hence encoded as transients 0, such that $n = 1$ and $u(\varepsilon \to 0) = 1/3$. Moreover, for thresholds ε larger than the largest distance between two samples, all points are merged into one single recurrence domain 1. However, the transient state 0 must always be present for reasons of consistency but is not realized in this case, such that $\boldsymbol{P}$ becomes the 2-dimensional unit matrix. This leads to $u(\varepsilon \to \infty) = 1/2$.

Application. – Optimal recurrence grammar encoding permits the detection of temporal segments in neurophysiological time series that best fit to recurrence domains [16]. The NRD serves as a measure of complexity in further evaluation. Recent studies [18,19] have provided evidence that the complexity of electroencephalographic data measured under general anaesthesia reflects the level of consciousness of subjects. In order to evaluate the optimal recurrence grammar encoding in this context, we investigate spatially distributed local field potentials (LFP) measured in ferret visual cortex under anaesthesia (cf. [20] for all details). Briefly, we performed electrophysiological recordings in primary visual cortex in one adolescent female ferret in a dark room while the animal was head-fixed, first while awake and later when anesthetized. Electrophysiology was conducted with acute invasive insertion of 32-channel probes ($50\,\mu$m with contact spacing along the z-axis and the reference electrode located on the same shank 0.5 mm above the top recording site). Unfiltered signals were amplified, then digitized at a rate of 20 kHz and finally downsampled to a rate of 100 Hz. For anesthetized recordings, general anesthesia was maintained with isoflurane (0.5%, 0.75% or 1.0%) and continuous infusion of xylazine. At least 15 minutes elapsed after changing anesthetic concentration prior to starting a new recording, exceeding the duration required in our setup for the LFP to stabilize at the new anesthetic concentration. All three levels of anesthesia used in this study corresponded to a lack of behavioral responses. We have extracted a large set of trials, each lasting 10 s, for each condition. According to previous studies, anaesthesia strongly affects neural activity in the α-frequency band from 8 Hz to 12 Hz [20,21]. Consequently, we bandpass-filtered the data in the α-band and extracted instantaneous power by a Gabor filter at each electrode. This yields a 32-dimensional signal that serves as the input signal of our recurrence analysis. In addition, to test for linearity we have phase-randomized all time series in order to generate surrogate data which we analyzed identically to the original data.

Results. – In order to validate the new optimization procedure, we consider again the standard Lorenz system with the same numerical settings as in [13]. For each ε, we compute the recurrence plot, its symbolic sequence and the corresponding transition probabilities between symbols (see external Supplementary Material[1], sect. 2). The results are presented in fig. 2. The upper panel of fig. 2(a) displays the time series of the model coordinates x_1, x_2 (blue and green, respectively) and x_3 (red) of the Lorenz trajectory $\boldsymbol{x}_t$. The Markov utility function, eq. (5), in fig. 2(b) assumes its maximum at $\varepsilon^* = 1.7$ (compare with $\varepsilon^* = 1.9$ in [13]). Using this threshold for the optimal recurrence grammar encoding of the trajectory $\boldsymbol{x}_t$ yields the color-coded symbolic dynamics s' depicted in the bottom panel of fig. 2(a). We observe a 5-state Markov chain with 4 recurrence domains. The segmentation into the two Lorenz wings as recurrence domains is clearly visible validating the optimal estimation of the system's recurrence structure.

Nonlinear dynamical systems exhibit nontrivial temporal recurrence structures. To further evaluate the proposed method we compare the recurrence structure of the Lorenz system with recurrence structures in linear systems. To this end, we create two kinds of linear stochastic surrogates through shuffling and phase randomization. Both preserve the statistical distribution of the data but destroy any nonlinear dependences of the original time series, although only the latter retains the signal's linear autocorrelation structure [22].

Figure 2(c), (d) shows the Lorenz surrogate data. Obviously, the maximum of the utility function for the phase-randomized surrogates (fig. 2(d), solid line) is substantially suppressed compared to that of the original time series. The optimal recurrence grammar for these surrogates yields a Markov chain with 6 recurrence domains (as compared to 4 in the original data), as seen in fig. 2(c). These correspond to smooth peaks of the time series that are spuriously detected as saddle nodes [14]. For comparison, fig. 2(d) also gives the utility function for time-shuffled surrogates (dotted line in panel (d)) mimicking white noise (compare with the external Supplementary Material, sect. 1). The function $u(\varepsilon)$ approaches $u(\varepsilon^*) = 1/2$ as its maximum that corresponds to a single recurrence domain. This result illustrates that a system without any essential recurrence structure is described by a regular sequence of only one recurrence domain, as in the case of white noise (see external Supplementary Material, sect. 1).

Next, fig. 3(a) shows the instantaneous α-power of the original single trial LFP from an animal anesthetized with 0.5% isoflurane in 32 channels distributed over space together with the optimal recurrence sequence. We observe a sequence of spatially distributed activity states in the 32-dimensional data (top panel) and the gained optimal sequence of symbolic states (bottom panel) that shows good accordance to the experimental data (see similar representations in the external Supplementary Material, sect. 4). Hence, the spatiotemporal sequences are

---

[1]We provide external Supplementary Material at `https://www.researchgate.net/publication/301525222_GrabenEA_EPLsupp`.
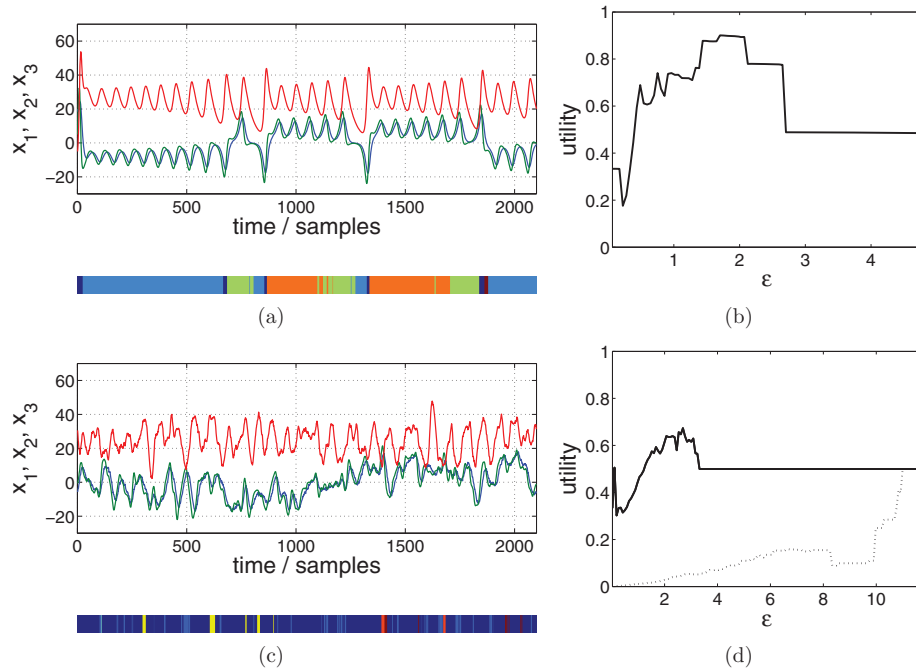
(a)

(b)

(c)

(d)

Fig. 2: (Color online) (a), (b): optimal recurrence grammar partition of the Lorenz attractor [17]. (a) Time series $\boldsymbol{x}_t$ (upper panel) and optimal encoding s' (color bar beneath). Color codes: x_1: blue; x_2: green; and x_3: red. (b) Markov utility function $u(\varepsilon)$ (eq. (5)). (c), (d): optimal recurrence grammar partition of Lorenz surrogates. (c) Phase-randomized surrogates $\boldsymbol{x}'_t$ (upper panel) and corresponding optimal encoding s' (color bar beneath). (d) Markov utility function $u(\varepsilon)$ (eq. (5)) for phase-randomized (solid) and shuffled surrogates (dotted).
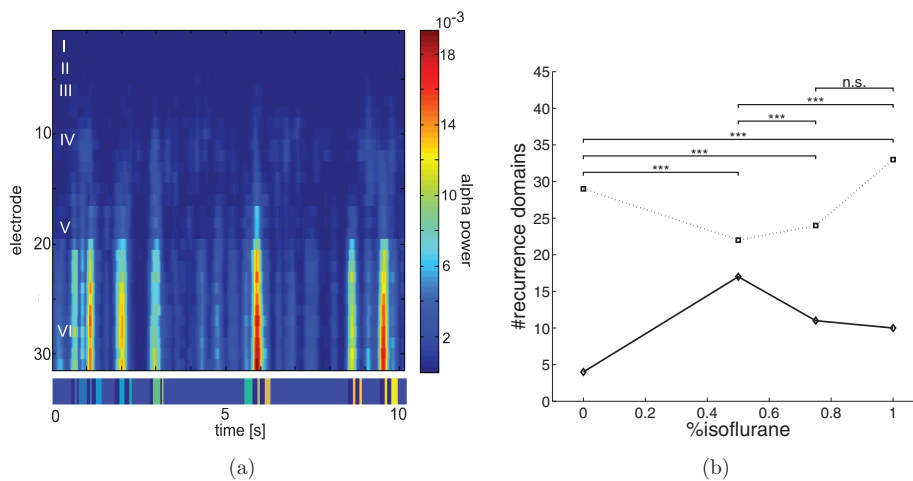

(a)

(b)

Fig. 3: (Color online) Neural activity in the α-frequency band for various isoflurane concentrations. (a) Time-resolved spatial distribution of power in a single trial under isoflurane concentration of 0.5% (top) where electrodes from 1 to 6 denote the supragranular cortical layers I–III, electrodes 7–14 the granular layer IV and 15–32 the infragranular layers V–VI. The corresponding color-coded optimal symbolic sequence s' shows good accordance to the distributed activity (bottom). The trial is selected to have the medium number of recurrence domains for 0.5% concentration, cf. panel on the right. (b) Median number of recurrence domains subject to the isoflurane concentration for the original data (solid line) and phase-randomized surrogates (dotted line). The data sets contains 132 (0%), 193 (0.5%), 180 (0.75%) and 176 (1.0%) trials, together with pairwise statistical comparisons using a Kolmogorov-Smirnov test.

identified well using the optimal recurrence grammar encoding. We determined the distribution of NRD over all trials for each concentration of anesthetic (illustrated in the external Supplementary Material, sects. 2 and 4).

These distributions are bimodal and thus not normal. Hence, fig. 3(b) displays medians for both the original and surrogate time series. Firstly, we observe a strong dependence of NRD on the concentration of anesthetic

which is significant for almost every pairwise comparison using a Kolmogorov-Smirnov test ($p < 10^{-6}$) between all concentrations in the original data set, except for the difference between 0.75% and 1.0% ($p > 0.05$). Moreover, results from original and phase-randomized data are groupwise significantly different (Kruskall-Wallis test, $p < 0.05$) demonstrating that the optimal recurrence grammar encoding clearly distinguishes nonlinear from linear dynamics. The awake state (0%) exhibits fewer recurrence domains than the anesthetized states with isoflurane concentrations larger than 0.5%. This reflects less different recurrent states in the brain, repetitions of fewer states and perhaps a higher degree of transient dynamics in the awake state compared to the anesthetized states. After the increase of NRD from 0% to 0.5% reflecting less regular signals, the significant decrease of NRD from 0.5% to 0.75% and 1.0% indicates a progression towards higher regularity, *i.e.*, ordering under deeper anaesthesia. The change of regularity may result from changes of intrinsic noise levels in neural populations. In fact, increased concentrations of anaesthetic agents diminish the activity input from subcortical to cortical structures leading to more regular neural activity [23]. This final decrease of regularity is in line with the well-established classification of EEG during anaesthesia: EEG becomes more synchronous with deeper anaesthesia before reaching levels associated with burst suppression. Our result in the α-frequency range shows a maximum in NRD indicating a new non-homogeneous relation between anaesthetic concentration and regularity.

Discussion. – The present work proposes a novel optimal estimate of recurrence structures that allows for the characterization of nonlinear temporal structures in multivariate time series by the number of recurrence domains. In contrast to previously suggested heuristics for the optimization of recurrence thresholds [2,12–14], our approach explicitly models the system's recurrence structure as transitions in a Markov chain between metastable states and transient regimes. We obtain a small number of recurrence domains with relatively large dwells that can be related to transitive subnetworks in a recurrence network close to the percolation transition [15]. The proposed method also shows similarities with other methods for finding a system's *generating partition* through partitioning homoclinic connections [24], as metastable states could be segmented along heteroclinic connections. We leave this interesting issue for future research.

$$* * *$$

REFERENCES

[1] POINCARÉ H., *Acta Math.*, **13** (1890) 1.
[2] MARWAN N., ROMANO M. C., THIEL M. and KURTHS J., *Phys. Rep.*, **438** (2007) 237.
[3] ECKMANN J.-P., KAMPHORST S. O. and RUELLE D., *Europhys. Lett.*, **4** (1987) 973.
[4] CASDAGLI M. C., *Physica D*, **108** (1997) 12.
[5] SOUZA E. G., VIANA R. L. and LOPES S. R., *Phys. Rev. E*, **78** (2008) 066206.
[6] JAVORKA M., TRUNKVALTEROVA Z., TONHAJZEROVA I., LAZAROVA Z., JAVORKOVA J. and JAVORKA K., *Clin. Physiol. Funct. Imaging*, **28** (2008) 326.
[7] LI X., SLEIGH J. W., VOSS L. J. and OUYANG G., *Neurosci. Lett.*, **424** (2007) 47.
[8] KARAGIANNI S. and KYRTSOU C., *Stud. Nonlinear Dyn. Econom.*, **15** (2011) 1558.
[9] OYA S., AIHARA K. and HIRATA Y., *New J. Phys.*, **16** (2014) 115015.
[10] HUNICZ J., GECA M., RYSAK A., LITAK G. and KORDOS P., *J. Vibroeng.*, **15** (2013) 1093.
[11] CARRIÓN A., MIRALLES R. and LARA G., *Ultrasonics*, **54** (2014) 1904.
[12] MARWAN N., *Int. J. Bifurcat. Chaos*, **21** (2011) 1003.
[13] BEIM GRABEN P. and HUTT A., *Phys. Rev. Lett.*, **110** (2013) 154101.
[14] BEIM GRABEN P. and HUTT A., *Proc. R. Soc. London, Ser. A*, **373** (2015) 20140089.
[15] DONGES J. F., HEITZIG J., DONNER R. V. and KURTHS J., *Phys. Rev. E*, **85** (2012) 046105.
[16] TOSIC T., SELLERS K. K., FRÖHLICH F., FEDOTENKOVA M., BEIM GRABEN P. and HUTT A., *Front. Syst. Neurosci.*, **9** (2016) 184.
[17] LORENZ E. N., *J. Atmos. Sci.*, **20** (1963) 130.
[18] SCHARTNER M., SETH A., NOIRHOMME Q., BOLY M., BRUNO M., LAUREYS S. and BARRETT A., *PLoS ONE*, **10** (2015) e0133532.
[19] ZHANG X.-S., ROY R. and JENSEN E., *IEEE Trans. Biomed. Eng.*, **48** (2001) 1424.
[20] SELLERS K., BENNETT D., HUTT A. and FROHLICH F., *J. Neurophysiol.*, **110** (2013) 2739.
[21] PURDON P. L., PIERCE E. T., MUKAMEL E. A., PRERAU M. J., WALSH J. L., WONG K. F., SALAZAR-GOMEZ A. F., HARRELL P. G., SAMPSON A. L., CIMENSER A., CHING S., KOPELL N. J., TAVARES-STOECKEL C., HABEEB K., MERHAR R. and BROWN E. N., *Proc. Natl. Acad. Sci. U.S.A.*, **110** (2012) E1142.
[22] PRICHARD D. and THEILER J., *Phys. Rev. Lett.*, **73** (1994) 951.
[23] HUDETZ A. G. and ALKIRE M. G., *Thalamocortical system and anesthetic-induced consciousness*, in *Neuroscientific Foundations of Anesthesiology*, edited by MASHOUR G. A. and LYDIC R. (Oxford University Press) 2011, pp. 45–61.
[24] GRASSBERGER P. and KANTZ H., *Phys. Lett. A*, **113** (1985) 235.