Journal of Neuroscience Methods 226 (2014) 33–41

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Computational Neuroscience

Application of adaptive nonlinear Granger causality: Disclosing network changes before and after absence seizure onset in a genetic rat model

Marina V. Sysoeva[a,*], Evgenia Sitnikova[b], Ilya V. Sysoev[a], Boris P. Bezruchko[a,c], Gilles van Luijtelaar[d]

[a] Saratov State University, 83, Astrakhanskaya str., Saratov, Russia
[b] Institute of Higher Nervous Activity and Neurophysiology RAS, 5A, Butlerova str., Moscow, Russia
[c] Saratov Branch of Institute of Radioengineering and Electronics of RAS, 38, Zelenaya str., Saratov, Russia
[d] Biological Psychology, Donders Centre for Cognition, Radboud University Nijmegen, Montessorilaan, 3, 6525 HR, Nijmegen, The Netherlands

HIGHLIGHTS

- Cortico-thalamic network associations were analyzed in rats with absence epilepsy.
- The outcomes of linear and adaptive nonlinear Granger causality were compared.
- Adaptive nonlinear measures were more sensitive to preictal changes of associations.
- Nonlinear interdependencies increased 1-1.5 s prior to seizure onset.

ARTICLE INFO

Article history:
Received 17 December 2013
Received in revised form 22 January 2014
Accepted 23 January 2014

Keywords:
WAG/Rij rats
Network analysis
Absence epilepsy
Adaptive modeling
Granger causality

ABSTRACT

Background: Advanced methods of signal analysis of the preictal and ictal activity dynamics characterizing absence epilepsy in humans with absences and in genetic animal models have revealed new and unknown electroencephalographic characteristics, that has led to new insights and theories.
New method: Taking into account that some network associations can be considered as nonlinear, an adaptive nonlinear Granger causality approach was developed and applied to analyze cortico-cortical, cortico-thalamic and intrathalamic network interactions from local field potentials (LFPs). The outcomes of adaptive nonlinear models, constructed based on the properties of electroencephalographic signal and on statistical criteria to optimize the number of coefficients in the models, were compared with the outcomes of linear Granger causality.
Results: The nonlinear adaptive method showed statistically significant preictal changes in Granger causality in almost all pairs of channels, as well as ictal changes in cortico-cortical, cortico-thalamic and intrathalamic networks. Current results suggest rearrangement of interactions in the thalamo-cortical network accompanied the transition from preictal to ictal phase.
Comparison with existing method(s): The linear method revealed no preictal and less ictal changes in causality.
Conclusions: Achieved results suggest that this proposed adaptive nonlinear method is more sensitive than the linear one to dynamics of network properties. Since changes in coupling were found before the seizure-related increase of LFP signal amplitude and also based on some additional tests it seems likely that they were not spurious and could not result from signal to noise ratio change.

1. Introduction

Absence epilepsy is non-convulsive generalized epilepsy of unknown etiology. Clinically, absence seizures appear as an abrupt and brief impairment of consciousness (absence), when ongoing activity is interrupted, responsiveness is decreased, and mental functioning is impaired. Electroencephalographically, absence seizures are manifested as paroxysmal electrical

* Corresponding author. Tel.: +7 9093367208.
E-mail addresses: bobrichek@mail.ru (M.V. Sysoeva), jenia-s@mail.ru (E. Sitnikova), ivssci@gmail.com (I.V. Sysoev), bezruchkobp@gmail.com (B.P. Bezruchko), g.vanluijtelaar@donders.ru.nl (G. van Luijtelaar).

0165-0270/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jneumeth.2014.01.028

activity consisting of generalized 3–4 Hz spike–wave discharges (SWD) (Panayiotopoulos, 2001).

Spike-and-wave paroxysms appear spontaneously in rat strains with a genetic predisposition to absence epilepsy, such as GAERS (Genetic Absence Epilepsy Rats from Strasbourg) and WAG/Rij (Wistar Albino Glaxo from Rijswijk). These two rat strains, GAERS and WAG/Rij, have been validated as reliable genetic animal models of human absence epilepsy, and they are widely used in basic research toward mechanisms involved in the pathogenesis of this disease (Vergnes et al., 1987; Marescaux et al., 1992; Coenen and van Luijtelaar, 2003; Depaulis, 2006).

In WAG/Rij rat model, later also in GAERS, it was found that a specific cortical area in the somatosensory region initiates spontaneous SWD (Meeren et al., 2002; Polack et al., 2007), and that cortical mechanisms effectively control and drive widespread cortico-cortical and cortico-thalamic networks during absence seizures (van Luijtelaar and Sitnikova, 2006; David et al., 2008; Lüttjohann et al., 2012). The transition from the preictal to the ictal phase is characterized by changes in associations within the cortico-thalamo-cortical neuronal network. Previously we examined the spatiotemporal synchronization of the thalamo-cortical system in WAG/Rij rats by means of EEG coherence (Sitnikova and van Luijtelaar, 2006), and also the strength and directionality of cortico-thalamic relationships by means of Granger causality (Sitnikova et al., 2008). These analyses confirmed that especially the onset of SWD seems a rather abrupt process in persons with absence epilepsy (Panayiotopoulos, 2001). However, other signal analytical approaches demonstrated the existence of preictal activity immediately before the onset of SWD both in persons with absence epilepsy (Inouye et al., 1994; Gupta et al., 2011) and in rats (Meeren et al., 2002; van Luijtelaar et al., 2011a; Lüttjohann et al., 2012; Lüttjohann and van Luijtelaar, 2012), perhaps in agreement with preictal neuronal firing in the deep cortical layers (Polack et al., 2007). Prediction of seizure activity from local field potentials (LFP) or surface EEG is a challenging problem that encourages specialists in physics and mathematics to develop new approaches of EEG data analysis (Mormann et al., 2007) that might be extremely relevant from a clinical perspective.

In comparison with traditional methods of network analysis, such as cross-correlation, coherence, phase synchronization, Granger causality may detect weak or hidden coupling, which not necessarily lead to synchronization, and defines next to changes in coupling strength also changes in the direction of coupling within a network. Granger causality was developed originally to detect the presence and direction of coupling between two systems (Granger, 1969). It takes into account the past state of one time series in order to predict the present state of the second time series. In his original paper, Granger (1969) used only linear predictive (autoregressive) models; new nonlinear models were successfully applied more recently (Bezruchko and Smirnov, 2010; Wang, 2007). Granger causality method is based on the idea to use competing models with optimal predictive abilities. The choice of the appropriate parameters in the model (parameterization) is important: even in linear Granger causality, the choice of the dimension of the model (the number of points in the past that are used for prediction) has a large impact on its predictive abilities. In nonlinear models, the type and number of nonlinear terms is also important. The proposed method actually depends on specific parameters of data, first on frequencies. As it was shown by Sysoeva et al. (2012), changing method parameters linked with frequency alters sensitivity and specificity of the method (number of real couplings that were not detected and number of false positive findings respectively). In this work used values of these parameters correspond to best specificity with as good sensitivity as possible.

If we need to construct vector time series from scalar observables (this usually refers to complex neurophysiologic signals, like EEG or LPF data), embedding takes place. Therefore parameters of embedding also become parameters of Granger causality method, for instance to determine the optimal number of time lags (Packard et al., 1980). Finally, when a prediction model is constructed, prediction length (the distance in time between time series point to be predicted and time series point(s) used for prediction) becomes another parameter of the model.

Wrong parameterization may cause false results: too simple models lead to missing couplings (Chen et al., 2004; Smirnov and Bezruchko, 2012; Sysoev et al., 2010; Sysoeva et al., 2012), too complex and universal models lead to unreliable estimates of model coefficients and often - to spurious causality. However it is possible to improve the methods by adapting the model structure to the experimental data: this implies that properties of the experimental data are used to set parameters of the model, in which case less coefficients can be used. Finally, a shorter length of time series can be used for model construction, given the opportunity to have a higher temporal resolution while applying Granger causality in a moving time window. Here we developed and applied this new approach called adaptive Granger causality to LFP data recorded in vivo by means of intracranial electrodes implanted in the cortex and thalamus in WAG/Rij rats. In particular, we investigated the dynamics of cortico-cortical, cortico-thalamic and thalamo-thalamic network interactions at the transition from preictal to ictal phase and compared results of linear and adaptive nonlinear estimation of Granger causality.

2. Methods

2.1. Animals and LFP data acquisition

Experiments were performed in five male 11–12 month old WAG/Rij rats. The recordings were done at the Department of Biological Psychology, Radboud University Nijmegen in accordance with the European Communities Council Directive (86/609/EEC). Experiments were approved by the Ethical Committee on Animal Experimentation of Radboud University Nijmegen. Distress and suffering of animals were minimal.

Rats were implanted, under complete inhalation anesthesia (isoflurane), with two standard tripolar electrode sets (Plastics One MS-333/2-A, Plastic Products, Roanoke, USA). There were stainless steel insulated wire electrodes with non-insulated tip (diameter 0.2 mm). Two epidural electrodes were located epidurally over the frontal (AP 2; L 2.5) and occipital (AP −7; L 6) cortical areas, skull flat. Two depth electrodes were implanted in the ventroposteromedial nucleus of thalamus (VPM, AP −3.5; L 2.5; H 7.2) and in the rostral pole of the reticular thalamic nucleus (RTN, AP −1.5; L 2.2; H 7.2). All coordinates are given in mm relative to bregma (Paxinos and Watson, 2006). Recording electrodes were implanted unilaterally at the right hemisphere. Ground and reference electrodes were placed symmetrically over both sides of the cerebellum. Electrodes were permanently attached to the rat's skull with dental cement.

After the surgery, animals were allowed to recover during at least ten days. During this recovery period, animals received post surgery care and their weight was monitored. Upon completion of the recording sessions, rats were deeply anesthetized with overdose of sodium pentobarbital (200 mg/kg i.p.) and their brains were stained with Nissl. Electrode positioning was verified using the atlas of the rat brain (Paxinos and Watson, 2006).

Recordings were performed in freely moving rats in a Faraday cage. Each recording session lasted from 5 to 7 h during the dark period of the day–night cycle. LFP signals were fed into a multi-channel differential amplifier, filtered between 1 and 200 Hz, digitized with 1024 samples/s/per channel (CODAS software) and stored on hard disk.

SWD were detected off-line in the frontal channel using the criteria of van Luijtelaar and Coenen (1986). Briefly, SWD appeared as a train of stereotypic repetitive 7–10 Hz spikes-and-waves with high amplitude (that exceeded the background more than three times) and lasted longer than 1 s.

2.2. Granger causality

The method considers two signals, X and Y that were recorded simultaneously from two brain areas and used for constructing two time series, $\{x_n\}_{n=1}^{N}$ and $\{y_n\}_{n=1}^{N}$, correspondingly at the nth time point. The study of causal (driving) interactions between X and Y with the aid of Granger causality includes three steps.

First, an univariate predictive model was constructed based on the one-channel raw data $\{x_n\}_{n=1}^{N}$ in the form of model map (1).

$$x'_{n+\tau} = f(x_n, x_{n-l}, \ldots, x_{n-(D_s-1)l}) \tag{1}$$

where $\vec{x}_n = (x_n, x_{n-l}, \ldots, x_{n-(D_s-1)l})$ is a state vector as defined by means of the method of delays (Packard et al., 1980), which is a classical approach to transpose time series in phase space, i.e. to obtain the high-dimensional state vector $\{\vec{x}_n\}_{n=1}^{N-(D-1)l}$ from the scalar time series $\{x_n\}_{n=1}^{N}$ for each time point. In this method, all components of the vector series were obtained from the same observable time series by shifting it back in time interval with duration of $l\,D_s$ times. So l is the time delay (or lag), and D_s is embedding dimension that is actually the number of components in a state vector (Kougioumtzis, 1996). Since the reconstructed vector $\{\vec{x}_n\}_{n=1}^{N-(D-1)l}$ was further used for modeling, its dimension D_s represented the dimension of univariate model (1), where $x'_{n+\tau}$ is the predicted value corresponding to the measured value $x_{n+\tau}$, τ is the length of prediction interval (prediction length), i.e., the time lag between the last point used for vector reconstruction and predicted point. Model coefficients were selected using the least square estimates (Legendre, 1805), i.e., by minimizing the squared prediction error (2), that measures the difference between predicted and observed values, $x'_{n+\tau}$ and $x_{n+\tau}$:

$$\varepsilon_s^2 = \frac{1}{N'\sigma_x^2} \sum_{n=(D_s-1)l}^{N-\tau} (x'_{n+\tau} - x_{n+\tau})^2 \to \min \tag{2}$$

where σ_x^2 is the dispersion of time series $\{x_n\}_{n=1}^{N}$, N' is an efficient length of time series, it is calculated as $N' = N - \tau - (D_s - 1)l$.

Second, the bivariate model (3) was constructed based on the both time series $\{x_n\}_{n=1}^{N}$ and $\{y_n\}_{n=1}^{N}$:

$$x''_{n+\tau} = g(x_n, x_{n-l}, \ldots, x_{n-(D_s-1)l}, y_n, \ldots, y_{n-(D_a-1)l}) \tag{3}$$

where D_a is dimension of the state vector $\vec{y}_n = (y_n, y_{n-l}, \ldots, y_{n-(D_s-1)l})$ reconstructed from the scalar time series $\{y_n\}_{n=1}^{N}$ in (3). So the total dimension of the bivariate model can be computed as $D_j = D_s + D_a$, and its prediction error is ε_j^2.

Third, the value of *prediction improvement PI* was computed with (4), and it is considered as the most important measurable characteristic of the adaptive Granger causality method

$$PI = 1 - \frac{\varepsilon_j^2}{\varepsilon_s^2} \tag{4}$$

The situation when $\varepsilon_j^2 = \varepsilon_s^2$ suggests that the data from the second EEG channel Y, i.e. time series $\{y_n\}_{n=1}^{N}$, do not improve prediction of $\{x_n\}_{n=1}^{N}$. In other words, Y does not drive X. The situation when $\varepsilon_s^2 > 0$ and $\varepsilon_j^2 \to 0$, providing $PI \to 0$ suggests that the data of the second time series $\{y_n\}_{n=1}^{N}$ significantly improves prediction of $\{x_n\}_{n=1}^{N}$, suggesting that Y drives X.

The outcomes of Granger causality method depends on model parameters, such as basis function (Chen et al., 2004; Marinazzo et al., 2006), polynomial order (Kornilov and Sysoev, 2013; Sysoev et al., 2010), time lag and prediction length (Sysoeva et al., 2012). Eventually, the model function and its parameters are crucially important for prediction quality and practical application of Granger causality method and should be taken into account to achieve reliable results. The univariate/bivariate model itself and its parameters should be accurately adjusted in order to avoid misleading results in EEG-derived data. The adjustment of model parameters will be explained in details in the next subsection.

2.3. Application of Granger causality to LFP data

Previously (Sitnikova et al., 2008) a linear model (5) was used with time lag and prediction length equal to one and optimized only parameters of model dimensions and, where optimization was performed based on the prediction error saturation criterion (Bezruchko and Smirnov, 2010).

$$\begin{aligned} x'_{n+1} &= c_0^s + \sum_{i=1}^{D_s} c_i^s x_{n-(i-1)}, \\ x''_{n+1} &= c_0^j + \sum_{i=1}^{D_s} c_i^j x_{n-(i-1)} + \sum_{i=D_s+1}^{D_s+D_a} c_i^j y_{n-(i-D_s-1)}, \end{aligned} \tag{5}$$

where c_i^s were empirically fitted coefficients in the univariate model and c_i^j were coefficients of the bivariate model.

Our further investigations were aimed in developing an adaptive nonlinear model (Sysoeva and Sysoev, 2012) that was specifically adjusted to the rodent EEG experimental data (Sysoeva et al., 2013a) (see (6)).

$$\begin{aligned} x'_{n+1} &= \sum_{k=0}^{P} \sum_{q=1}^{C_{D_s+k}^k} c_i^s \prod_{m=1}^{D_s} x_{n-(m-1)l}^{w_{k,m}^s} + c_{Z_s+1} x_{n-l_T}, \forall k \\ &= 0, \ldots, P \sum_{m=1}^{D_s} w_{k,m}^s = k \end{aligned} \tag{6a}$$

$$\begin{aligned} x''_{n+1} &= \sum_{k=0}^{P} \sum_{q=1}^{C_{D_s+D_a+k}^k} c_i^j \prod_{m=1}^{D_s} x_{n-(m-1)l}^{w_{k,m}^j} \times \prod_{m=1}^{D_a} y_{n-(m-1)l}^{w_{k,(m+D_s)}^j} \\ &\quad + c_{Z_j+1} x_{n-l_T} + c_{Z_j+2} y_{n-l_T} \quad \forall k = 0, \ldots, P \sum_{m=1}^{D_s+D_a} w_{k,m}^j = k, \end{aligned} \tag{6b}$$

where $Z_s = (P + D_s)!/(P!\,D_s!)$ is the number of coefficients in the univariate model (6a), $Z_j = (P + D_s + D_a)!/(P!\,(D_s + D_a)!)$ is the number of coefficients in the bivariate model (6b), $C_{D_s+k}^k$ is the number of combinations varied from $D_s + k$ to k, P is the polynomial order, l_T is the additional lag that took into account the value of the experimental data delayed from the predicted time point with a period of T.

In the adaptive nonlinear model (6), approximating functions f and g were generalized polynomials; a similar approach with the same polynomial functions was used earlier (Chen et al., 2004). The prediction length value in the model (6) was set to $\tau = T/4$ as it was proposed in Sysoeva et al. (2012), where T is the duration of one characteristic period in a signal. The chosen value of τ provided the best compromise between *sensitivity* of the presented method (capability to detect the actual coupling) and *specificity* (minimum of false positive decisions about coupling). The value

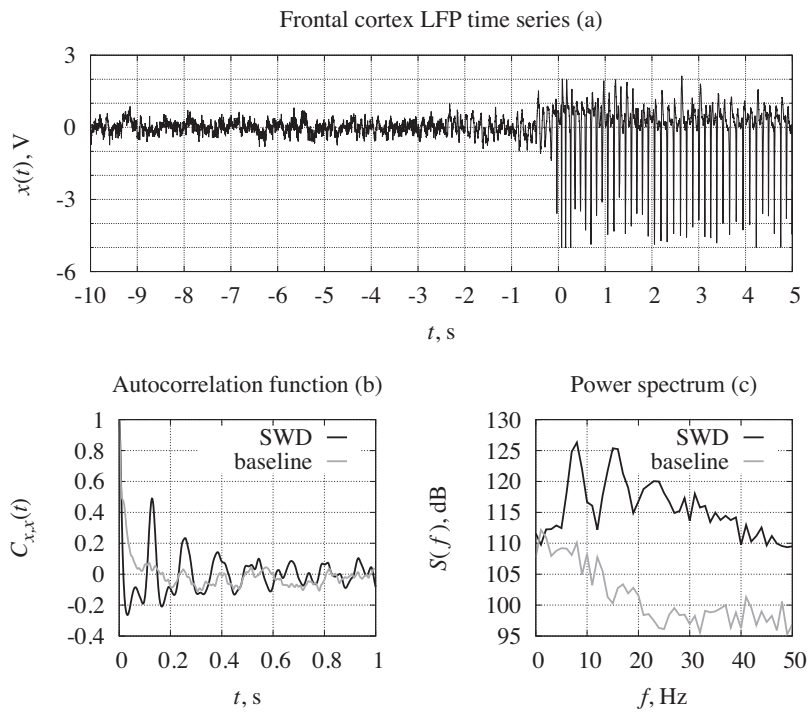

Fig. 1. An example of epileptic spike–wave discharges (SWD) in WAG/Rij rat: (a) LFP (time series) as recorded in frontal cortex before the seizure (10 s) and SWD (5 s); (b) autocorrelation function, calculated from 10 s before of seizure (baseline) and 5 s of seizure (SWD); (c) power spectrum, calculated from the same intervals as the autocorrelation functions in control periods (at least 10 s before seizure onset, baseline) and of seizure (SWD) (Bartlett, 1948).

T for $\{x_n\}_{n=1}^{N}$ was determined as the first maximum in its autocorrelation function (Fig. 1b) or in power spectrum (Fig. 1c). LFP data before and during SWD (Fig. 1a) were non-stationary (more details about non-stationarity in electrical brain activity can be found in Dikanev et al. (2005)). Despite the fact that the onset of SWD caused remarkable changes in LFP, autocorrelation functions of LFP periods before and during SWD (Fig. 1b) contained similar regular components with characteristic period T. It was obvious though that this regular component was more pronounced in the full blown SWD as compared with the period before SWD onset. As far as the length of the period T before and during SWD was almost the same, the value of T as measured during SWD was used to assess the values of l and τ in pre-seizure interval, therefore the model developed in Sysoeva and Sysoev (2012) could be applied for Granger causality analysis before and during SWD.

Polynomial order P and model dimension D_s must be carefully selected and verified in order to minimize the risk of underestimation (when the model is too simple and unable to reproduce the dynamics of the LFP) and overestimation (when the number of coefficients is too high and their values cannot be reliably defined based on the LFP data). The quality of prediction of models (5) and (6) is based on the prediction least-squared errors ε_s^2 and ε_j^2. If the computation accuracy is high, than any increase in the number of coefficients decreases the errors ε_s^2 and ε_j^2. Therefore, additional statistical information criteria, such as Bayesian information criterion (BIC) or Akaike criterion, should be used in order to determine the optimal values of P and D_s (Wang, 2007; Akaike, 1974; Schwarz, 1978).

Previously (Sitnikova et al., 2008), a prediction error saturation criterion was used to define the values of D_s and D_a. In this criterion, the threshold value for the difference between least-square errors of models with different number of coefficients is introduced, e.g. let us denote it as $\Delta_{\max}\varepsilon_s^2$ for ε_s^2. This threshold value was necessary to define the last coefficient in the model: the model did not require additional coefficients when the difference between errors for model with Z_s coefficients and the model with $Z_s + 1$ coefficients was less than $\Delta_{\max}\varepsilon_s^2$. In this case, the number of coefficients Z_s in the model was considered as optimal. This approach might result in overcomplicated models, mainly because of subjectivity in selecting the value $\Delta_{\max}\varepsilon_s^2$. In order to tackle this problem in the present study, the optimal value of the number of coefficients was determined automatically.

Methods of automatic finding of the number of coefficients are usually based on optimization of a target function. This function is a combination of the prediction error ε_s^2 and penalty term that depends on number of model coefficients, Z_s. According to this approach, medium-range values of Z_s were optimal, because small models were rejected due to too the high value of the error ε_s^2, and large models were rejected due to too large penalty term. In our previous study (Sysoeva and Sysoev, 2012), the polynomial order P and model dimension D_s were determined based on the Bayesian information criterion (BIC, also known as Schwarz criterion) (Schwarz, 1978) with target function (7).

$$S = \frac{N'}{2}\ln\left(\varepsilon_s^2\right) + Z_s\frac{\ln(N')}{2} \tag{7}$$

In the general nonlinear model, as proposed by Chen et al. (2004), the number of model coefficients would be so high that the coefficient values could not be determined reliably from the LFP data, even though the polynomial order P and dimension D_s were selected with BIC. In our study (Sysoeva and Sysoev, 2012), it was shown that non-uniform embedding procedure (see Judd and Mees, 1998) could be used to reduce the number of coefficients without significant effect on model quality. This approach was used in the current study, and it was introduced as additional linear terms with second lag l_T in both equations of system (6). The optimal value of l_T was chosen, according to formula (Eq. (8)) (Sysoeva and Sysoev, 2012).

$$\tau + l_T = T \tag{8}$$

Table 1
Basic features of linear and adaptive nonlinear models applied in Granger causality.

	Linear model (5)	Adaptive nonlinear model (6)
Nonlinearity	Not considered	Considered
Time scales	Not considered. Lag (l) and prediction length (τ) were used, and both were equal to 1 point $\tau = l = 1$.	Considered. Three time scales were used and derived from the LFP signals: 1 $\tau = T/4$, 2 $l = T/10$, 3 $l_T = T - \tau$.
Optimization of the number of coefficients	Prediction error saturation criterion with empirically defined threshold (subjective factor)	Bayesian information criterion (BIC) (Schwarz, 1978) for the fully automatic optimization (subjective factor was excluded)
Sensitivity and *specificity*	Generally high sensitivity for quasi-stationary states, insufficient specificity – many spurious coupling detections	Sufficient high sensitivity with good specificity
Detection of coupling *direction*	Problems with coupling direction detection due to spurious effects	Ability to reliably detect coupling direction

The value of lag l was optimized based on the fact that, in systems with characteristic period T, the best specificity could be obtained with the lag varying between $T/12$ and $T/3$ (Kornilov et al., 2013), and the values $l = 1$ and $l = nT/2$ (where n is an integer number) must be avoided, because of a very high probability of false positive results. Using two different lags l and l_T however may lead to problem of considering twice the same point. Indeed, to predict point n the same point $n - T$ can be taken into account first setting $l_T = T - \tau$ and second with setting $l = (T - \tau)/(D_s - 1)$. Since we used dimension $D_s = 6$ for some models, we had to use small enough value of lag l to avoid this problem. Therefore we set $l = T/10$.

Table 1 summarizes some basic features of linear and adaptive nonlinear models applied for the analysis of LFP signals.

In our experimental data, the start and ending moments of each SWD were defined in LFP recordings. Baseline (preictal) period included 10 s before the seizure onset. The first 5 s of seizure activity was identified as the ictal period.

In order to estimate time-dependent changes of coupling characteristics between different brain areas and especially with the onset of SWD, we used principles of time-variant Granger causality (Hesse et al., 2003), and moving window that was shifted in time by 0.1 s. The prediction improvement PI (4) in our experimental data was calculated in a moving window of 0.5 s (512 points) in both linear and adaptive nonlinear models.

Resulting dependencies of PI on time were averaged across all seizures in each animal, matching start moments of seizures. SWD that lasted less than 5 s were ignored, with the numbers of seizures included into analysis for different animals being 34, 94, 10, 22, and 58 respectively. Then for each averaged dependency $PI(t)$ the background level PI_{bg} was established as an average PI over 7 s time interval (baseline period, from 10 to 3 s before the SWD start). Very similar results could be achieved if one uses another interval as a baseline, e.g. from 10 to 7 s before the SWD onset, with using longer interval provides us little bit more significant differences in terms of PI between baseline and preictal/ictal phase. Then normalized dependencies were calculated as $PI_0(t) = PI(t) - PI_{bg}$. The value of $PI_0 = 0$ means the baseline level, and positive values of PI mean larger coupling than in baseline and negative – lower coupling.

3. Results

Linear (5) and adaptive nonlinear (6) models were applied in the framework of Granger causality approach for the analysis of cortico-cortical, thalamo-thalamic and cortico-thalamic interactions in LFP recordings during baseline period, before and after SWD onset (preictal and ictal phases).

Fig. 2 displays the dynamics of linear and nonlinear Granger causality (namely, normalized to baseline level of prediction improvement PI_0) during 10 s interval before the onset of SWD and during the first 5 s of SWD. The moment of SWD onset was determined as the first spike in SWD train.

In linear Granger causality, the mean amplitude of PI (i.e., coupling strength) did not differ from zero in any of the channel pairs in either direction before the onset of SWD and remained constant throughout the whole 10 s interval before SWD onset (baseline and preictal period, Fig. 2). The maximum amplitude of PI was found only ~0.5 s after the onset of SWD, when the moving time window covered the fully developed seizure. In nonlinear Granger causality, changes in coupling were more complex: in the majority of channel pairs, PI amplitude increased ~1.5–3 s before the onset of SWD, reached a maximum ~0.5 s prior to SWD onset and dropped at the beginning of ictal phase. Therefore, adaptive nonlinear Granger causality was sensitive to changes in cortical-cortical, cortico-thalamo-cortical and intrathalamic network coupling during the preictal period, suggesting that this is due to seizure precursor activity. As a remark, we need to say that the same increase could be observed in other physiological conditions (such as low vigilance state), and it might not always lead to seizures.

In order to statistically analyze time dynamics of cortico-thalamo-cortical associations with the aid of adaptive nonlinear Granger causality (Fig. 3), amplitudes of PI were plotted in a period between preictal phase (3 s prior to SWD onset) and ictal phase (3 s after SWD onset) and Student's t tests aiming at establishing differences from zero were carried out. In order to correct for an increased chance of false positives and to reduce the chance of getting type I errors, only clusters of minimal 3 subsequent points were considered as significant (Maris and Oostenveld, 2007). The amplitude of PI during baseline at 3 s prior to SWD onset did not differ from zero in all investigated pairs. A statistically significant increase from zero started first in thalamic to frontal cortex pairs (VPM → FC, RTN → FC) and from OC → FC at 2.2 s before the onset of SWD. From 1.2 s prior to SWD onset, the FC started to drive both thalamic channels, the OC got further involved via the VPM, and the OC influenced the RTN. At 0.7–0.4 s preceding SWD onset, the RTN replied the OC, the OC replied the VPM, the FC the OC, while intrathalamic pairs were among the slowest to increase PI amplitude.

The onset of SWDs was often accompanied by an increase in the amplitude of PI, as it was established with the linear Granger causality (Fig. 2). The largest changes were found in the extent that the frontal cortex increased its influence on VPM and RTN, while also the influence of the VPM on the RTN increased. Other increases in linear causality seemed smaller. Two kinds of changes were noticed with the nonlinear method, as could be seen in Fig. 2. In

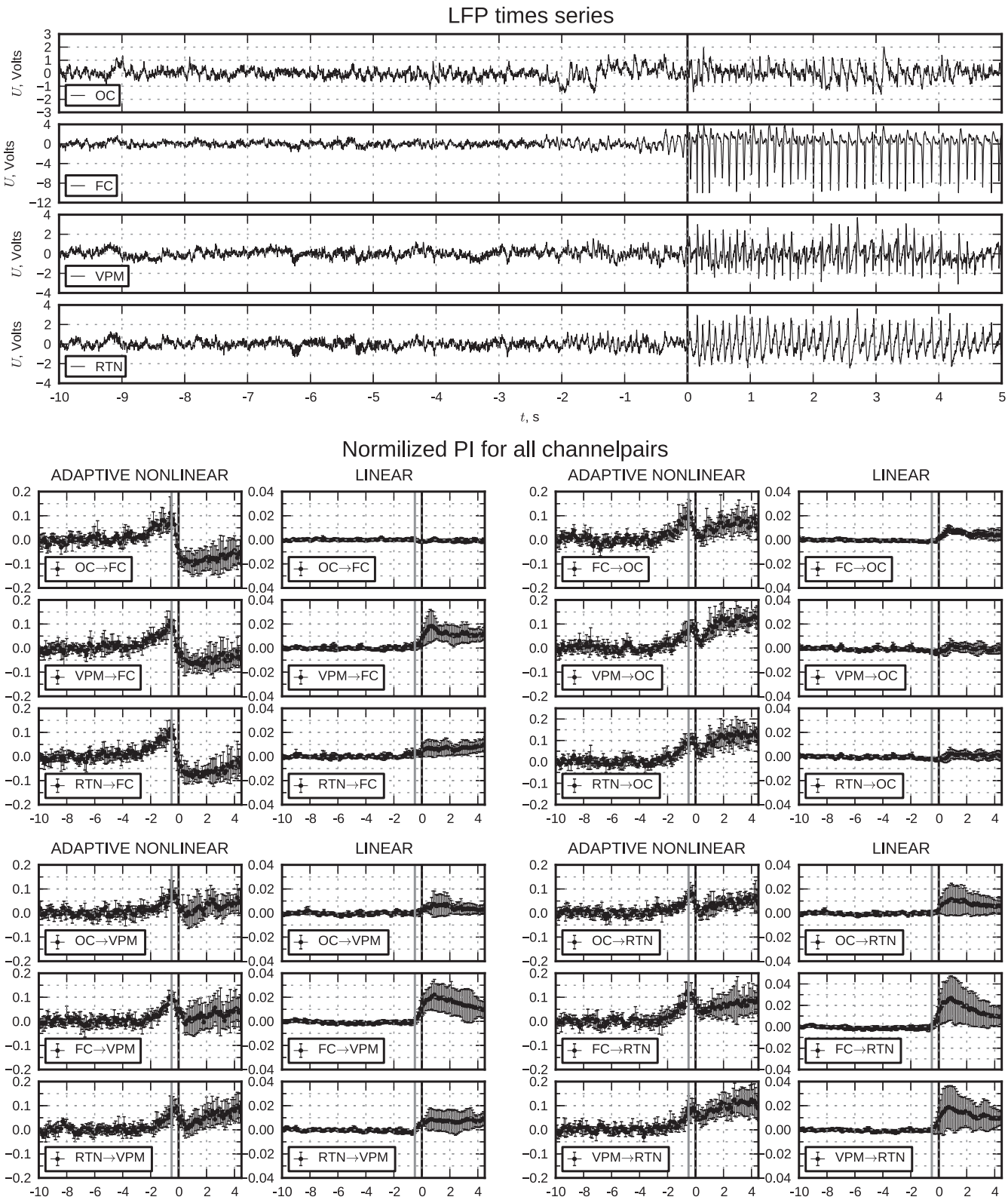

Fig. 2. (Top) LFP time series for all channels from 10 s before seizure onset till 5 s after it, i.e. the exact analyzed interval. The moment of the onset of SWD corresponds to time point zero. (Bottom) Dynamics of linear and adaptive nonlinear Granger causalities as measured in the period around the onset of SWD. *Y*-axis: amplitude of normalized *PI*, i.e. PI_0. Prediction improvement (*PI*) was averaged per rat (5 rats, S.E.M.'s are plotted): the number of SWD varied from 33 and 111 per rat. The plots were sorted by the driven structure. Black vertical line indicates the seizure onset, gray vertical line indicates the length of moving window, in which Granger analysis was performed.

the pairs where the frontal cortex was the structure being driven (OC → FC, VPM → FC, RTN → FC) the value of PI dropped to baseline level at the onset of SWD or became negative and lower than baseline. For other driven channels such as OC and RTN the PI first grew in the preictal phase, then dropped at the beginning of ictal phase and slowly increased during the first 3 s of ictal phase. The influence on the VPM remained low during the ictal phase. Such a picture could be explained like two different processes in the brain

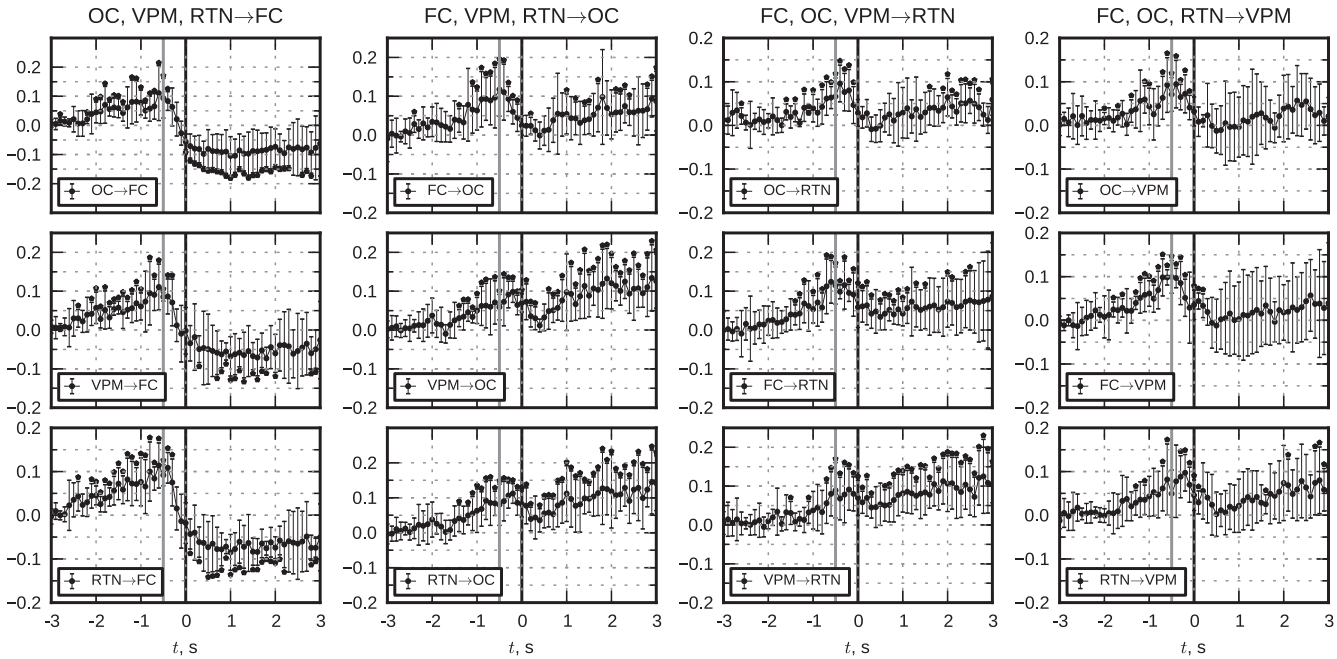

Fig. 3. Results of statistical analysis of adaptive nonlinear Granger causalities during preictal and ictal phases (3 s before and 3 s after SWD onset). *Y*-axis: amplitude of normalized *PI* (i.e. PI_0), averaged across subjects. Errorbars show the 95% confidence interval, calculated, considering 5 *PI* values for different subjects for each time point. Asterisks indicate significant differences from zero with *p*-value <0.05.

being activated. One process is responsible for seizure initiation and manifests itself as a gradual increase of coupling during preictal phase. This process impacts to all channels. The second process is responsible for seizure maintenance and manifests itself as increase in coupling during ictal phase, following relative decrease in coupling at the beginning of ictal phase. This second process is evident in pair of channels where OC and RTN are driven structures, leading to the overlap of two processes.

4. Discussion

The main result of the present study was that the comparison between linear and nonlinear adaptive estimators of Granger causality showed more and larger changes for the nonlinear variant. Second, the adaptive nonlinear Granger causality disclosed seizure-related changes of interactions between almost all thalamic and cortical pairs of electrodes before the onset of SWD. Finally, some of the interrelations remained increased during SWD.

The outcomes of the current study demonstrated that linear Granger causality was constantly low during preictal phase and started to increase when SWD were visually recognized in the LFP. This confirms and extends our previous findings (Sitnikova et al., 2008) that Granger causality based on a linear model is able to detect and quantify changes in coupling strength and direction between different cortical regions, between cortex and thalamus and within the thalamus during spontaneous mild seizures. However, a linear model was not sufficient to determine preictal changes. Preictal changes in local field potentials in WAG/Rij rats were previously found with wavelet analysis and other methods of time–frequency analysis (van Luijtelaar et al., 2011b; Lüttjohann et al., 2012), as well as in their coupling with nonlinear association analyses (Meeren et al., 2002; Lüttjohann and van Luijtelaar, 2012), demonstrating that changes in network interdependencies were associated with SWD and immediately preceded the onset of SWDs. Indeed, SWD do not arise suddenly, instead, they are locally initiated, and develop from increasing associations within and between cortical layers and subsequently subcortical regions, all before a "generalized" seizure becomes apparent (van Luijtelaar et al., 2011a).

Our calculations (Figs. 2 and 3) demonstrate that nonlinear influences to occipital cortex (FC, VPM, RTN → OC) and to RTN (FC, VPM, OC → RTN) during preictal period and during SWD were higher than in baseline. It is proposed that both the elevation of thalamo-cortical influences to occipital cortex and to RTN might be involved in seizure initiation and maintenance (generalization) processes. In all, the comparison between the two Granger methods shows a larger sensitivity for the nonlinear adaptive variant that is able to detect new and previously hidden network associations.

It needs to be remarked that no LFP electrodes were aimed at the focal zone in the somatosensory cortex. Therefore, these results should not be considered as the ultimate description of the cortico-cortical and cortical-thalamic networks. A much higher spatial resolution of the signals is necessary for that, which can be achieved with a combination of cortical grids and a sufficient set of thalamic depth electrodes. Moreover, the possibility that some of the causalities, such as the fronto-occipital one, may arise because of the presence of an active (focal) region which direct its influence to both the frontal and occipital cortex, cannot be excluded. Therefore, the outcomes of the present study should be considered as the development, description and evaluation of a new method and less of a complete description of network activities. In some cases, a decrease below baseline was seen following SWD onset. This can be explained either as an actual decrease in coupling, or as a result of significant simplification of the dynamics of the LFP in the frontal cortex at the seizure onset. The model that was well suited for revealing an increase in the amplitude of PI during the preictal period is no longer optimal for prediction of the subsequent points. Put into other words, the negative values may imply that the signal can explain its future by itself and that adding points from the other signal to predict subsequent points from the first one is actually not leading to a prediction improvement. It may point to reduced entropy in the signal at SWD onset, a point for further investigation.

The nonlinear adaptive Granger causality technique presented in this paper is based on an empirical model that takes properties of the considered signals into account. More specifically, the time scales of the original LFP signal were analyzed, also the dimension and nonlinearity of the model were carefully estimated based on BIC criterion. With a non-uniform embedding technique the number of model coefficients was additionally decreased, so the model can be fitted in the same short time window as the linear one previously used (Sitnikova et al., 2008). All these modifications gave us the opportunity to extract additional information from the same data. As a result, adaptive nonlinear Granger causality revealed changes of coupling in all channel pairs during ictal phase (SWD) and also during preictal phase, demonstrating a larger sensitivity of the new method to seizure-precursor activity and confirmed that SWDs do not arise suddenly.

It needs to be pointed that it is critical to perform every step of model's adaptation. First, inclusion of nonlinearity in the model improves sensitivity to preictal and ictal activity. Second, the incorporation of signal properties into the model (such as dominant frequencies of the LFP signals) prevents finding spurious couplings. Third, application of non-uniform embedding and BIC criterion for the optimal model dimension and polynomial order decreases the number of coefficients, and helps to solve the problem of their underestimation. This allows to use a small time window with good enough time resolution that appeared to be sufficient for detecting complex changes of network associations during preictal activity. Most likely, the absence of these adaptations significantly worsened prediction quality of model and led to sufficiently poorer results.

The achieved results cannot be explained by a change in the amplitude of the relevant signal with seizure onset, since changes in coupling often start 1 or even 1.5 s before the moving window includes the seizure activity. That is after the adaptations, the method became more sensitive to changes in associations, but not to changes in amplitude of the signals. This is in a good correspondence with results of Dikanev et al. (2005). Also some calculations performed for assembles of etalon nonlinear oscillators with different coupling strength suggested the independency of Granger causality on the amplitude of signal (Sysoeva et al., 2013b).

To sum, preictal SWD activities were associated with temporal increase of only nonlinear Granger causality in almost all network channel pairs, suggesting that nonlinear interdependencies increased through the entire network immediately prior to SWD onset.

Our current results indicated that the overall pattern of seizure-related dynamics of Granger causality differed in different pairs, suggesting that seizure-related interactions within these networks are characterized by prominent anisotropy.

Acknowledgements

We thank H. Krijnen, S. Menting-Hermeling, E. Willems-van Bree, J. Dederen for excellent technical assistance.

The work was supported by the RFBR (grants #12-02-00377, #13-04-00084, #14-02-00492).

References

Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control 1974;19:716–23.

Bartlett MS. Smoothing periodograms from time-series with continuous spectra. Nature 1948;161:686–7.

Bezruchko B, Smirnov D. Extracting knowledge from time series. Berlin: Springer; 2010.

Chen Y, Rangarajan G, Feng J, Ding M. Analyzing multiple nonlinear time series with extended Granger causality. Physics Letters A 2004;324(1):26–35.

Coenen AML, van Luijtelaar ELJM. Genetic animal models for absence epilepsy: a review of the WAG/Rij strain of rats. Behavior Genetics 2003;33:635–55.

David O, Guillemain I, Saillet S, Reyt S, Deransart C, Segebarth C, et al. Identifying neural drivers with functional MRI: an electrophysiological validation. PLoS Biology 2008;6:2683–97.

Depaulis A, van Luijtelaar G. Genetic models of absence epilepsy in the rat. In: Pitkanen A, Moshe S, Schwartzkroin P, editors. Animal models of seizures and epilepsy. San Diego: Elsevier Inc.; 2006. p. 223–48.

Dikanev T, Smirnov D, Wennberg R, Perez Velazquez JL, Bezruchko B. EEG nonstationarity during intracranially recorded seizures: statistical and dynamical analysis. Clinical Neurophysiology 2005;116:1796–807.

Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 1969;37(3):424–38.

Gupta D, Ossenblok P, van Luijtelaar G. Space–time network connectivity and cortical activations preceding spike wave discharges in human absence epilepsy: a MEG study. Medical & Biological Engineering & Computing 2011;49:555–65.

Hesse W, Molle E, Arnold M, Schack B. The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. Journal of Neuroscience Methods 2003;124:27–44.

Inouye T, Matsumoto Y, Shinosaki K, Iyama A, Toi S. Increases in the power spectral slope of background electroencephalogram just prior to asymmetric spike and wave complexes in epileptic patients. Neuroscience Letters 1994;173(May (1–2)):197–200.

Judd K, Mees A. Embedding as a modeling problem. Physica D 1998;120: 273–86.

Kornilov M, Golova T, Sysoev I. Choice of parameters for Granger causality method for systems with characteristic time scale. In: Proceedings of XIV Pan-Russian school-seminar "Physics and application of microwaves"; 2013. p. 57–60 [in Russian].

Kornilov M, Sysoev I. Influence of the choice of the model structure for working capacity of nonlinear Granger causality approach. Izv VUZov "Applied Nonlinear Dynamics" 2013;21(2):74–87 [in Russian].

Kougioumtzis D. State space reconstruction parameters in the analysis of chaotic time series – the role of the time window length. Physica D 1996;95(1): 13–28.

Legendre AM. Appendice sur la méthodes des moindres quarrés. Nouvelles méthodes pour la détermination des orbites des cométes. Paris: Firmin-Didot; 1805. p. 72–80 [in French].

Lüttjohann A, Stoffelen JM, van Luijtelaar G. Periictal network dynamics of spike–wave discharges: phase and spectral characteristics. Experimental Neurology 2012;239:235–47.

Lüttjohann A, van Luijtelaar G. The dynamics of cortico-thalamo-cortical interactions at the transition from pre-ictal to ictal LFPs in absence epilepsy. Neurobiology of Disease 2012;47:47–60.

Marescaux C, Vergnes M, Depaulis A. Genetic absence epilepsy in rats from Strasbourg – a review. Journal of Neural Transmission (Supplementum) 1992;35:37–69.

Marinazzo D, Pellicoro M, Stramaglia S. Nonlinear parametric model for Granger causality of time series. Physical Review E 2006;73:066216.

Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. Journal of Neuroscience Methods 2007;164(1):177–90.

Meeren HK, Pijn JP, van Luijtelaar EL, Coenen AM, Lopes da Silva FH. Cortical focus drives widespread corticothalamic networks during spontaneous absence seizures in rats. Journal of Neuroscience 2002;22:1480–95.

Mormann F, Andrzejak RG, Elger CE, Lehnertz K. Seizure prediction: the long and winding road. Brain 2007;130(2):14–333.

Packard N, Crutchfield J, Farmer J, Shaw R. Geometry from a time series. Physical Review Letters 1980;45:712–6.

Panayiotopoulos CP. Treatment of typical absence seizures and related epileptic syndromes. Paediatric Drugs 2001;3:379–403.

Paxinos G, Watson C. The rat brain in stereotaxic coordinates. 6th ed. San Diego: Academic Press; 2006.

Polack PO, Guillemain I, Hu E, Deransart C, Depaulis A, Charpier S. Deep layer somatosensory cortical neurons initiate spike-and-wave discharges in a genetic model of absence seizures. Journal of Neuroscience 2007;27:6590–9.

Schwarz G. Estimating the dimension of a model. The Annals of Statistics 1978;6(2):461–4.

Sitnikova E, Dikanev T, Smirnov D, Bezruchko B, van Luijtelaar G. Granger causality: cortico-thalamic interdependencies during absence seizures in WAG/Rij rats. Journal of Neuroscience Methods 2008;170(2):245–54.

Sitnikova E, van Luijtelaar G. Cortical and thalamic coherence during spike–wave seizures in WAG/Rij rats. Epilepsy Research 2006;71:159–80.

Smirnov D, Bezruchko B. Spurious causalities due to low temporal resolution: towards detection of bidirectional coupling from time series. Europhysics Letters 2012;100:10005.

Sysoev I, Karavaev A, Nakonechny P. Role of model nonlinearity for granger causality based coupling estimation for pathological tremor. Izv VUZov "Applied Nonlinear Dynamics" 2010;18(2):81–90 [in Russian].

Sysoeva M, Dikanev T, Sysoev I. Selecting time scales for empirical model construction. Izv VUZov "Applied Nonlinear Dynamics" 2012;20(2):54–62 [in Russian].

Sysoeva M, Dikanev T, Sysoev I, Bezruchko B. Analysis of coupling between rat electroencephalogram channels before and during epileptic seizure using predictive models. Vestnik of NNSU. Radiophysics 2013a;1(1):73–8 [in Russian].

Sysoeva M, Sysoev I. Mathematical modeling of encephalogram dynamics during epileptic seizure. Technical Physics Letters 2012;38(2):151–4.

Sysoeva MV, Sysoev IV, Sitnikova EYu. Application of adaptive Granger causality to revealing coupling structure with absence seizures. In: Proceedings of X international conference "Chaos-2013", Saratov 2013; 2013b. p. 42.

van Luijtelaar EL, Coenen AM. Two types of electrocortical paroxysms in an inbred strain of rats. Neuroscience Letters 1986;70:393–7.
van Luijtelaar G, Sitnikova E. Global focal aspects of absence epilepsy: the contribution of genetic models. Neuroscience and Biobehavioral Reviews 2006;30(7):983–1003.
van Luijtelaar G, Sitnikova E, Lëttjohann A. On the origin and suddenness of absences in genetic absence models. Clinical EEG and Neuroscience 2011a;42(2): 83–97.
van Luijtelaar G, Hramov A, Sitnikova E, Koronovskii A. Spike–wave discharges in WAG/Rij rats are preceded by delta and theta precursor activity in cortex and thalamus. Clinical Neurophysiology 2011b;122:687–95.
Vergnes M, Marescaux C, Depaulis A, Micheletti G, Warter JM. Spontaneous spike and wave discharges in thalamus and cortex in a rat model of genetic petit mal-like seizures. Experimental Neurology 1987;96:127–36.
Wang CW. Nonlinear phenomena research perspectives. New York: Nova Science Publishers; 2007. p. 7–53.