The development of statistical methods for the diagnosis of neurological diseases based on multiparameter analysis of brain activity

V A Yunusov¹ and S A Demin¹

¹Kazan Federal University, Institute of Physics, Kazan, 420008 Russia

E-mail: serge_demin@mail.ru

Abstract. The application of data science for the analysis of biomedical data presented as time series allows using new methods of analysis of living systems. Therefore, using these methods it could be possible to discover new diagnostic criteria for neurological diseases. In this paper, in the framework of Memory Functions Formalism, one of the approaches of statistical physics, we analyze recordings of signals of the human brain cortex in the sleep state of 19 patients with nocturnal frontal lobe epilepsy and 8 healthy subjects. We observe alterations of dynamic parameter behavior at the pathology and in healthy subjects. Furthermore, we reveal significant alterations in brain rhythms manifestations at statistical memory functions power spectra of patients with epilepsy and healthy subject. As a result, we show that the application of the statistical analysis of electroencephalogram recordings, after appropriate verification, can be helpful in the search of diagnostic criteria of nocturnal frontal lobe epilepsy.

1. Introduction
Nowadays data science is one of the most rapidly developing areas of computer science. Analysis of time series generated by complex systems of animate and inanimate nature plays a special role in data science. This is due to the active evolution in recording equipment technologies, the development and implementation of new software tools, and the increase in the computing power of the analysis equipment.

In data science, there is a constant search for new challenges, one of which is the search for diagnostic criteria for human diseases based on the use of information technologies. Biomedical data of a different nature are analyzed for the diagnosis of neurological diseases.

For the search for features in the signals recordings of human nervous system activity, various approaches of statistical analysis are used. In [1, 2], the authors studied the synchronization effects of correlation variables calculated for bioelectric activity signals of different brain regions. It allowed detecting differences in the degree of frequency–phase synchronization of the bioelectrical activity of the cerebral cortex. The signals of healthy subjects are characterized by well-defined effects of frequency–phase synchronization. On the contrary, brain activity signals of patients are characterized by a significant abnormality of frequency–phase synchronization.

In paper [3], the authors applied the methodology of the Flicker–Noise Spectroscopy to determine the difference in bioelectric activity response signals to flickering light stimuli of a group of healthy subjects and a patient with photosensitive epilepsy. As a result, they demonstrated the effectiveness of this method in identifying the individual features of the signals not only for a patient but also for
healthy subjects. This allows using Flicker–Noise Spectroscopy methodology for early diagnosis not only of photosensitive epilepsy but also for other neurological diseases.

Besides that, methods of machine learning are wildly applied to the detection and study of anomalies in the dynamics of human brain activity. In paper [4], the authors reviewed machine learning approaches for seizure detection. The authors concluded that machine learning classifiers are suitable for accurate seizure detection, but it’s crucial to select a suitable type of classifier for each particular case under study. Some types of classifiers can help to discover seizure localization and seizure type; the others can achieve higher predictive accuracy.

The development of methodologies for the analysis of human biomedical data allowed detecting and separating a large number of diseases into separate types. However, international criteria for the diagnosis of some diseases have not been established yet. One of the diseases of this type is nocturnal frontal lobe epilepsy (NFLE). NFLE is a syndrome characterized by the onset of seizures varying severity and duration during sleep state in the frontal lobe of the cerebral cortex.

In the study in which similar data were analyzed [5], the authors used threshold–dependent symbolic entropy and computed Normalized Corrected Shannon Entropy. The authors concluded that this method of time series analysis in comparison with other spectral methods of analysis more effectively distinguished healthy subjects from patients with NFLE. Therefore, threshold–dependent symbolic entropy methodology can be used as a mean for better diagnosis of this disease.

In this paper, we search for diagnostic criteria for NFLE within the framework of one of the data science methods – the Memory Functions Formalism (MFF). MFF is a theoretical approach that studies the dynamics of stochastic processes in complex systems.

2. Methods. Memory Functions Formalism

The temporal dynamics of an experimentally recorded parameter of a complex system of living nature can be represented as a discrete time series \( x_j \) of a variable \( X \):

\[
X = \{x(T), x(T + \tau), x(T + 2\tau),..., x(T + (N - 1)\tau)\},
\]

where \( T \) is the initial time from which the recording of the experimental parameter started, \((N - 1)\tau\) is the signal recording time, \( \tau = \Delta t \) is the sampling time step.

The average value of the dynamic variable \( \langle X \rangle \), fluctuations \( \delta x_j \) and absolute variance \( \sigma^2 \) can be represented as follows:

\[
\langle X \rangle = \frac{1}{N} \sum_{j=0}^{N-1} x(T + j\tau), \quad \delta x_j = x_j - \langle X \rangle, \quad \sigma^2 = \frac{1}{N} \sum_{j=0}^{N-1} \delta x_j^2.
\]

For a quantitative description of the dynamic properties of the living system under study (correlation dynamics), it is convenient to use the normalized time correlation function (TCF):

\[
a(t) = \frac{1}{(N-m)\sigma^2} \sum_{j=0}^{N-m-1} \delta x_j \delta x_{j+m} = \frac{1}{(N-m)\sigma^2} \sum_{j=0}^{N-m-1} \delta x(T + j\tau) \delta x(T + (j + m)\tau),
\]

where \( x_j \), \( x_{j+m} \) are the values of variable \( X \) on steps \( j \), \( j + m \) correspondingly, \( \delta x_j, \delta x_{j+m} \) are fluctuations of values \( x_j, x_{j+m} \), \( \sigma^2 \) is the absolute variance of the variable \( X \).

Using the technique of the Zwanzig–Mori projection operators [6, 7] introduced in nonequilibrium statistical physics allows to obtain a chain of finite difference equations of non–Markov type [8, 9] for the initial \( a(t) \) and higher-order memory functions \( M_i(t) (i=1,2,...,n) \):
\[
\frac{\Delta a(t)}{\Delta t} = \lambda_t a(t) - \tau N \sum_{j=0}^{m-1} M_j(t) a(t - j \tau), \ldots, \tag{3}
\]

\[
\frac{\Delta M_{m-1}(t)}{\Delta t} = \lambda_m M_{m-1}(t) - \tau N \sum_{j=0}^{m-1} M_j(t) M_{m-1}(t - j \tau),
\]

where \(\lambda_t\) are parameters that form the spectrum of eigenvalues of the Liouville quasi–operator \(\hat{L}\), \(\Lambda_j\) are relaxation parameters:

\[
\hat{\lambda}_n = i \frac{\langle W_{n-1} \hat{L} W_{n-1} \rangle}{\langle W_{n-1}^2 \rangle}, \quad \Lambda_n = \frac{\langle W_n^2 \rangle}{\langle W_{n-1}^2 \rangle}. \tag{4}
\]

Dynamic orthogonal variables \(W_n\) in equation (4) are obtained using the Gram–Schmidt orthogonalization procedure:

\[
\langle W_n, W_m \rangle = \delta_{n,m} \langle W_n^2 \rangle.
\]

where \(\delta_{n,m}\) is the Kronecker symbol.

In earlier papers [8, 9], to quantify the effects of statistical memory, the authors proposed a frequency dependence of the non–Markov parameter:

\[
\varepsilon_i(\nu) = \left( \mu_{i-1}(\nu) \right)^{\frac{1}{2}}, \tag{5}
\]

where the frequency characteristics of \(\mu_i(\nu)\) power spectra are determined through the Fourier images of the memory functions \(M_j(t)\):

\[
\mu_0(\nu) = \left| \Delta t \sum_{j=0}^{N-1} a(t_j) \cos 2\pi \nu t_j \right|^2, \quad \ldots, \mu_i(\nu) = \left| \Delta t \sum_{j=0}^{N-1} M_j(t) \cos 2\pi \nu t_j \right|^2. \tag{6}
\]

The non–Markov parameter \(\varepsilon = \varepsilon_i(0)\) (i.e. the value of the statistical memory measure at zero frequency) allows distinguishing Markov processes (with short-range memory) and non–Markov processes (with long-range memory). An analysis of the values of the non–Markov parameter calculated for various biomedical data suggests that it also can indicate the pathological state of the complex system (e.g. human body as a whole or human brain).

In this work, we implemented above–mentioned expressions in the MATLAB environment to analyze stochastic dynamics variables and spectral parameters of experimental data under study.

3. Experimental data

Experimental data were obtained at the Sleep Disorders Center of the Ospedale Maggiore of Parma, Italy. It consists of interictal multichannel EEG (electroencephalography) recordings from 8 healthy subjects with no evidence of neurological diseases and 19 patients with NFLE. The recordings were made during sleep state from electrodes \(A_1, F_4, F_8\), using electrode \(C_4\) as a reference electrode, by the International electrode placement system “10–20” (shown in figure 1) and published in Physionet database [10]. The average duration of EEG–signals recording of the subjects was about 8 hours.
Figure 1. The scheme of the “10–20” International electrode placement system.

4. Search for the diagnostic criteria using autocorrelation analysis
As a result of the autocorrelation analysis of the multichannel EEG recordings for two groups of subjects, the values of non–Markov parameters were calculated. Using these values, a “meaningful” electrode, which corresponds to the area of the brain cortex, where pathological activity manifested the most in NFLE, was identified. Identifying the meaningful electrode before the main part of the analysis is necessary to narrow the range of search for possible diagnostic criteria, and to determine those signals, the dynamics of which differs the most for two groups of subjects. For the subsequent comparative analysis of the obtained results of autocorrelation analysis, EEG signals of people received from the electrode $A_1$ were selected, with the values of non–Markov parameters $\epsilon(0)$ close to the average values in their group. In the control group, it is subject number 4, in the group of people with NFLE, it is patient number 14.

The peculiarity of the MMF is that the dynamic and spectral characteristics of the signal under study are revealed from the spectrum of various parameters: phase portraits of orthogonal dynamic variables, power spectra of memory functions, relaxation and kinetic parameters. As a result of the autocorrelation analysis, the features of the dynamics of the studied parameters for the EEG signals of two groups of subjects were established.

Phase portraits of orthogonal dynamic variables for both subjects were built on the same scale for a visual comparison of the manifested structural features and the form of the variables' values distribution. The structure of phase portraits for a healthy person (figures 2a, 2c, 2e) is ordered; stratification is not observed for the signal from an electrode. At the pathology (figures 2b, d, f) there is a significant stratification in the structure of phase portraits for signals from all electrodes, which is manifested the most in the signal from electrode $A_1$. 
Figure 2. Phase portraits of orthogonal dynamic variables for different electrodes for a healthy subject (a, c, e) and a patient with nocturnal frontal lobe epilepsy (b, d, f).

Besides that, different types of the spectral behavior of EEG signals were found in healthy subjects and patients with epilepsy. On the graphs of the power spectra of the normalized TCF and the statistical memory functions (figures 3, 4), a manifestation of different rhythms of the brain cortex activity for the healthy subject and the patient with epilepsy is shown.

Figure 3. Power spectra of the time correlation function and statistical memory functions of EEG signals for a healthy subject. Arrows denote frequencies with the most intense activity.
Figure 4. Power spectra of the time correlation function and statistical memory functions of EEG signals for a patient with nocturnal frontal lobe epilepsy. Arrows denote frequencies with the most intense activity.

For the control group subject δ- and θ-rhythms are manifested the most (1.5 Hz and 8 Hz respectively), corresponding to a state of restful natural sleep. For the patient with NFLE δ–rhythm (0.3 Hz) was manifested, which corresponds to a state of natural sleep. Besides, β– and γ–rhythms (in range of 35–90 Hz) are most prominent in power spectra of statistical memory functions, which are naturally manifested in a state of active wakefulness and high concentration.

Also, at the pathology the presence of bursts of brain activity can be observed, indicating the existence of “overlapping” periodic brain activities. The manifestation of such cyclic characteristics may indicate the presence of a disease, and, therefore, serve as criteria for the diagnosis of this pathology.

During the search for diagnostic criteria for NFLE, graphs for all representatives of the control group and patients with epilepsy were analyzed. The mean values of the scatter of graphs parameters and mean values of frequencies of the peaks with the greatest intensity were determined and are presented in Table 1.

Table 1. The mean values of the studied parameters of the graphs for the control group subjects and patients with NFLE.

|         | \( \bar{W}^{\min}_{0} (\tau) \) | \( \bar{W}^{\max}_{0} (\tau) \) | \( \bar{W}^{\min}_{1} \) | \( \bar{W}^{\max}_{1} \) | \( \bar{v}_{\mu_1}^{\min} \) (Hz) | \( \bar{v}_{\mu_1}^{\max} \) (Hz) | \( \bar{v}_{\mu_2}^{\max} \) (Hz) | \( \bar{v}_{\mu_2}^{\min} \) (Hz) |
|---------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Control | 5               | 9               | 4               | 6               | 8               | 10              | 9               | 8               |
| Patients| 143             | 328             | 17              | 89              | 0.7             | 46              | 35              | 76              |

The choice of the above–mentioned patient with NFLE and healthy subject from their groups for the comparative analysis is because the values of graphs parameters for these subjects are close to the mean values of these parameters for the respective groups. Therefore, this patient and this healthy subject are considered typical representatives of their groups.
5. Results and conclusions
In this paper in the framework of Memory Functions Formalism for the search of diagnostic criteria for NFLE autocorrelation analysis of EEG recording was performed. The studied parameters of the system include non–Markov parameters, phase portraits of orthogonal dynamic variables, and power spectra of statistical memory functions. The comparing of this variety of parameters for the healthy subject and the patient with the pathology leads to the following conclusions.

During the autocorrelation analysis, we discovered an area of the brain cortex in which pathological activity was manifested the most; we revealed significant alterations in dynamic parameter behavior at the pathology in comparison with a healthy subject.

Besides that, we discovered $\beta$– and $\gamma$–rhythms of brain functioning, which are not typical for the studied state of patients. That partially coincides with the results of other works [11], in which, during the analysis of nocturnal interictal EEG recordings, the manifestation of $\beta$– and $\theta$–rhythms of brain activity was found in patients. Also, in subjects with nocturnal frontal lobe epilepsy, we revealed significant manifestations of abnormal bursts of brain activity. The results obtained during the research, after appropriate verification, could be used to determine the diagnostic criteria for nocturnal frontal lobe epilepsy.

Further prospects for the search of diagnostic criteria for this disease are related to the application of other methods of time series analysis: cross-correlation analysis [12, 13], Flicker–Noise Spectroscopy [2, 3], fractal structure analysis [14–16].

Acknowledgments
The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University. This work is supported by the Russian Science Foundation (project no. 20-12-00105).

References
[1] Panischev O Y, Demin S A and Bhattacharya J 2010 Physica A 389 4958
[2] Timashev S F, Polyakov Y S, Yulmetyev R M, Demin S A, Panischev O Y, Shimojo S and Bhattacharya J 2010 Las. Phys. 20 604
[3] Timashev S F, Polyakov Y S, Yulmetyev R M, Demin S A and Panischev O Y 2009 Las. Phys. 19 836
[4] Siddiqui M K, Morales-Menendez R, Huang X and Hussain N 2020 Brain Inform. 7 5
[5] Hussain L, Shah S A, Aziz W, Bukhari S N H, Lone K J and Chaudhary Q A 2020 Waves Random Complex Media 1 1
[6] Zwanzig R 1961 Phys. Rev. 124 983
[7] Mori H 1965 Prog. Theor. Phys. 34 399
[8] Demin S A, Panischev O Y, Demina N Y and Yunusov V A 2020 CEUR Workshop Proc. 2667 110
[9] Yulmetyev R, Demin S, Emelyanova N, Gafarov F and Hänggi P 2003 Physica A 319 432
[10] Terzano M G, Parrino L, Sherieri A, Chervin R and Chokroverty S 2001 Sleep Med. 2 537
[11] Derry C P, Harvey A S and Walker M C 2009 Sleep 32 1637
[12] Panischev O Y, Demin S A, Kaplan A Y and Varaksina N Y 2013 Biomed. Eng. 47 153
[13] Ghosh D, Dutta S and Chakraborty 2014 Chaos Solition. Fract. 67 1
[14] Andreev A O, Demina N Y, Demin S A, Nefedev Y A and Churkin K O 2016 NPCS 19 271
[15] Lahmiri S 2018 Biomed. Signal Process Control 40 312
[16] Lahmiri S 2018 Physica A 490 378