

# Complexity of EEG Signals in Schizophrenia Syndromes

I.E. Kutepov<sup>1</sup>, V.A. Krysko<sup>1</sup>, A.V. Krysko<sup>1</sup>, S.P. Pavlov<sup>1</sup>, M.V. Zigalov<sup>1</sup>, I.V. Papkova<sup>1</sup>, O.A. Saltykova<sup>1</sup>,  
T.Y. Yaroshenko<sup>1</sup>, E.Y. Krylova<sup>2</sup>, T.V. Yakovleva<sup>1</sup>, V.V. Dobriyan<sup>1</sup>, N.P. Erofeev<sup>1</sup>  
ilyakutepov@yandex.ru

<sup>1</sup>Yuri Gagarin State Technical University of Saratov, Saratov, Russia;

<sup>2</sup>Saratov State University, Saratov, Russia

*In the present study, 45 patients with schizophrenia syndromes and 39 healthy subjects are studied with electroencephalogram (EEG) signals. The study groups were of different genders. For each of the two groups, the signals were analyzed using 16 EEG channels. Multiscale entropy, Lempel-Ziv complexity and Lyapunov exponent were used to study the chaotic signals. The data were compared for two groups of subjects. Entropy was compared for each of the 16 channels for all subjects. As a result, topographic images of brain areas were obtained, illustrating the entropy and complexity of Lempel-Ziv. Lempel-Ziv complexity was found to be more representative of the classification problem. The results will be useful for further development of EEG signal classification algorithms for machine learning. This study shows that EEG signals can be an effective tool for classifying participants with symptoms of schizophrenia and control group. It is suggested that this analysis may be an additional tool to help psychiatrists diagnose patients with schizophrenia.*

**Keywords:** entropy, chaos, EEG classification, schizophrenia, Lyapunov exponent.

## 1. Introduction

Schizophrenia is associated with disorders in the lobes and areas of the brain, which are responsible for information processing, temporary memory and executive functions [3]. The diagnosis of schizophrenic spectrum disorders and other psychotic disorders is challenging. The scientific community is constantly working to integrate the latest clinical and scientific advances in the field of psychiatry into diagnostic and statistical manuals. [5]

However, quantifying and evaluating abnormalities in the cerebral cortex can help to understand the mechanisms of such psychotic disorders. Recent advances in the area of analysis of complexity of time series provide insights into nonlinear electroencephalogram (EEG) signals. [1]. The complexity of time series can be investigated by using several measures, for instance, Approximate Entropy or Sample Entropy - SampEn. Traditional entropy-based algorithms quantitatively determine the regularity (ordering) of a time series. Entropy rises as the degree of irregularity increases and is maximum for completely random systems. However, an increase of entropy is not always associated with an increase of dynamic complexity. For example, randomized time series have a higher entropy than the original time series, since the process of generating of surrogate data reduces the correlation and worsens the information content of the original signal.

It is worth noting that many methods have been developed for estimating the complexity of time series based on entropy presented in the review [6], but preference is given to Multiscale Entropy - MSE. Multiscale entropy relies on sample entropy calculations at different scales: the MSE algorithm uses the SampEn algorithm to analyze time series that represent the system dynamics at various levels. Multiscale entropy has become the predominant method for quantifying the complexity of signals. This method has been successfully used in various fields of research, including biomedical time series [2]. The disadvantages of this method include: a discrete representation of a signal of continuous nature (significantly affects on the entropy estimate), signal length, presence of noise, selection of parameters (length of the analyzed sequence, cell size of the phase space).

Another common method for estimating of the complexity of EEG signals is the Lempel - Ziv complexity - LZC. This method is non-parametric, model-independent and easily calculated. In addition, it does not require long time series [12]. The LZC algorithm provides more reliable results for short signal segments, which is important in most experimental and clinical studies [4].

The oscillatory character by virtue of EEG is indicated of the hypothesis that EEG signals originate from a nonlinear dynamic system. Therefore, the unpredictability of the EEG can be considered as a phenomenon characterized by randomness. The essential property of chaotic dynamics is the so-called sensitive dependence on the initial conditions. This property can be quantified by calculating the first positive Lyapunov exponent (L1) in the system [8]. Studies [10] showed that patients with schizophrenia, the values of Lyapunov's senior exponent were lower in the left lower frontal and anterior temporal areas compared with the control group.

The purpose of this study is to compare the signal complexity estimates obtained by the MSE, LZC methods and the Lyapunov senior exponent. It is suggested that nonlinear EEG analysis can be a useful tool in the analysis of EEG data for studying the neurodynamics of the brain of patients with schizophrenia.

## 2. Methods.

### 2.1. Subject of research.

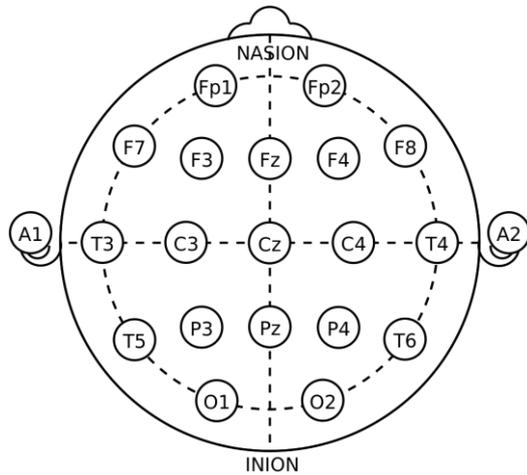
Two EEG data archives were analyzed for two groups of subjects [http://brain.bio.msu.ru/eeg\_schizophrenia.htm]. The subjects of the survey were adolescents who were tested by a psychiatrist and divided into two groups: healthy (n = 39) and with symptoms of schizophrenia (n = 45). Each file contains an EEG record for one subject. Each TXT file contains a column with EEG samples from 16 EEG channels, according to Fig.1. Signals were recorded by channels: 'F7', 'F3', 'F4', 'F8', 'T3', 'C3', 'Cz', 'C4', 'T4', 'T5', 'P3', 'Pz', 'P4', 'T6', 'O1', 'O2'. Each number in the column represented the EEG amplitude ( $\mu V$ ) on a separate sample. The first 7680 samples represent 1 channel, then 7680 - channel 2, etc. The sampling rate is 128 Hz, so 7680 samples correspond to 1 minute of EEG recording.

### 2.2. Multiscale Entropy

The entropy calculation method MSE was presented in [2].

For a given discrete time series  $\{x_1, \dots, x_i, \dots, x_N\}$ , the sequence is determined from the simplified time series  $\{y^{(\tau)}\}$  with respect to the scaling parameter  $\tau$ . The original time series is divided into non-overlapping windows with a length  $\tau$ , and then the values are averaged for each window. Thus, each element of the simplified time series is calculated by the formula

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, \quad 1 \leq j \leq N/\tau.$$



**Fig. 1.** Configuration diagram of electrodes on the surface of head

For the first scale, the time series  $\{y^{(1)}\}$  is equivalent to the original time series. The length of each time series corresponds to the length of the original time series divided by the scaling parameter  $\tau$ .

The calculation of the quantitative measure of entropy  $S_E$  for each simplified time series is made by the formula

$$S_E(m, r, N) = \ln \frac{\sum_{i=1}^{N-m} n_i^m}{\sum_{i=1}^{N-m} n_i^{m+1}},$$

where  $m$  – the increment of the length of the data vector,  $r$  – the cell size of the phase space (inaccuracy),  $n_i^m$  – the probability of repeating a sequence of data of a given length in the original data.

### 2.3. Lempel – Ziv complexity

Lempel and Ziv proposed a measure of the complexity of patterns for sequences of finite length [9]. Later, Kaspar and Shuster developed an algorithm for computing the LZC on a computer that determined the measure of complexity [7]. LZC calculates the number of new images, i.e. segments that are not consistently represented in all previous data. In this algorithm, the EEG signal  $\{x(n)\}$  is converted into a binary sequence  $\{s(n)\}$  by comparing with the average value of the signal  $m$ . After receiving a binary sequence, the corresponding measure of complexity  $c(n)$  is increased by one until a new sequence is detected. The process of searching for sequences is repeated until the last character of the time series is read. LZC is defined as

$$LZC = c(n)/b(n),$$

где  $b(n) = n/\log_2(n)$ .

### 2.4. Lyapunov exponent.

Lyapunov exponent give an estimate of the average exponential divergence or convergence of nearby trajectories in the phase space. Obtaining a positive value of the Lyapunov's exponent is characterized by dependence on the initial conditions and shows that the system of interest is chaotic.

To calculate the Lyapunov senior exponent. L1, a modified version of the Wulf algorithm was used [11]. Essentially, the tangent vectors to points on the reference path are approximated by difference vectors in the phase space. The Wulf algorithm is based on the fact that the time series is normalized to match the equilibrium state to zero. Next is the reconstruction of the trajectory in the phase space. After that, with some step for the coordinate vector of the reconstructed phase space, the indicator component is calculated. For each component calculation, the series is renormalized so that the initial discrepancy coincides for each next component. Then the procedure is repeated.

## 3. Results

For a comparative analysis of complexity (MSE, LZC, L1), the same EEG signals of two groups were studied. The average values for each of the EEG recording channels were calculated for the control group (norm) and patients with schizophrenia syndromes (sch.). Statistical analysis based on P-value was used to compare methods of signal complexity. P-value is the probability that the criterion value will be not less than the critical value, provided that the null hypothesis about the absence of differences between groups is true. If the P-value was less than 0.05, the difference between the mean values was considered significant. Thus, the method that will show the largest number of channels with  $P < 0.05$  will be considered the most characteristic.

The results of calculation of MSE, LZC, L1 presented in Tables 1-3 are consistent with the majority of previous works, showing that patients with schizophrenia are characterized by less complex neuropsychological measurements. The analysis showed that the most characteristic channels for MSE entropy assessment are: 'F3', 'F4', 'T6'.

**Table 1** Average MSE

Channel	Control group (Average + SD)	Sch group (Average + SD)	P- value
'F7'	0,9296 ± 0,1154	0,8859 ± 0,1495	0,1568
'F3'	0,9524 ± 0,1306	0,8867 ± 0,1267	0,0296
'F4'	0,9655 ± 0,1230	0,8866 ± 0,1221	0,0042
'F8'	0,9629 ± 0,1563	0,9035 ± 0,1385	0,1075
'T3'	0,9670 ± 0,1363	0,926 ± 0,1386	0,1705
'C3'	0,9604 ± 0,1047	0,9309 ± 0,1043	0,2085
'Cz'	0,9454 ± 0,1087	0,9165 ± 0,0866	0,2180
'C4'	0,9412 ± 0,1026	0,9212 ± 0,1025	0,4021
'T4'	0,9737 ± 0,1329	0,9346 ± 0,1256	0,1587
'T5'	0,9891 ± 0,1586	0,9608 ± 0,1354	0,3733
'P3'	0,9267 ± 0,1063	0,9151 ± 0,1117	0,6317
'Pz'	0,9044 ± 0,1029	0,8907 ± 0,1031	0,5801
'P4'	0,9161 ± 0,1090	0,8955 ± 0,1091	0,4415
'T6'	0,9922 ± 0,1446	0,9112 ± 0,1323	0,0094
'O1'	0,8822 ± 0,1130	0,9007 ± 0,1271	0,4969
'O2'	0,8747 ± 0,1086	0,8843 ± 0,1332	0,7071

The analysis of the data presented in Table 2 shows that the objects in the control group are characterized by a lower average value of LZC compared to objects with schizophrenia syndromes. The most indicative channels for LZC estimation are: 'F4', 'Cz', 'P3', 'O1', 'O2'.

It should be noted that the complexity of the Lempel-Ziva characterizes the chaotic nature of the signal with an inverse value in relation to the entropy and Lyapunov exponent, i.e. the higher the value of LZC the more regular the time series.

Table 3 presents average values of Lyapunov exponent L1 for two groups of research objects. The analysis of the average values of Lyapunov senior exponent for all study objects showed that the most typical results were obtained through the channel 'O1'.

Visualization of the data which was obtained in the Matlab software package, algorithms for constructing topographic images of the EEG signal spectra shown in Fig.2. Topographic images are constructed in accordance with the layout of the electrodes in Fig. 1. The channels used in the research are marked with dots. Values at intermediate points were interpolated by using a spherical spline.

**Table 2** Average LZC

Channel	Control group (Average + SD)	Sch group (Average + SD)	P- value
'F7'	0,0128 ± 0,0592	0,0284 ± 0,0813	0,3514
'F3'	0,1418 ± 0,0261	0,1542 ± 0,0365	0,0793

'F4'	0,1415 ± 0,0294	0,1597 ± 0,0406	0,0239
'F8'	0,1667 ± 0,0413	0,1862 ± 0,0459	0,0631
'T3'	0,1707 ± 0,0371	0,1766 ± 0,0540	0,5992
'C3'	0,1382 ± 0,0298	0,1517 ± 0,0391	0,1282
'Cz'	0,1317 ± 0,0292	0,1469 ± 0,0362	0,0497
'C4'	0,1370 ± 0,0306	0,1524 ± 0,0377	0,0576
'T4'	0,1722 ± 0,0544	0,1765 ± 0,0490	0,7377
'T5'	0,1668 ± 0,0433	0,1762 ± 0,0434	0,3813
'P3'	0,1261 ± 0,0322	0,1487 ± 0,0402	0,0085
'Pz'	0,1197 ± 0,0291	0,1353 ± 0,0387	0,0626
'P4'	0,1251 ± 0,0338	0,1427 ± 0,0395	0,0536
'T6'	0,1619 ± 0,0471	0,1601 ± 0,0464	0,8733
'O1'	0,1057 ± 0,0332	0,1338 ± 0,0412	0,0025
'O2'	0,1032 ± 0,0324	0,1280 ± 0,0416	0,0083

MSE(a)

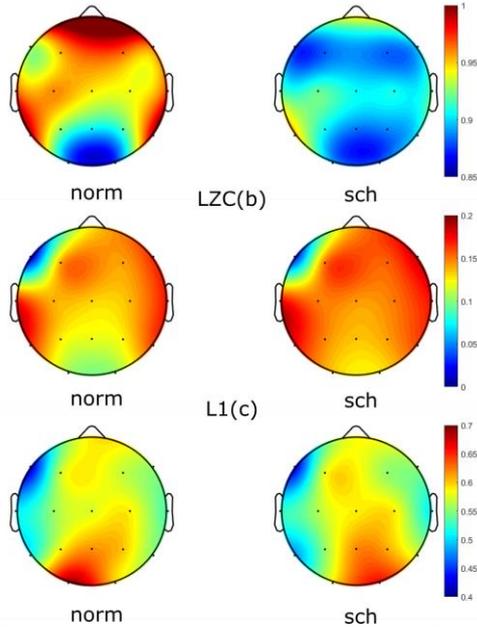


Fig. 2. Topographical representations of MSE(a), LZC(b) and L1(c)

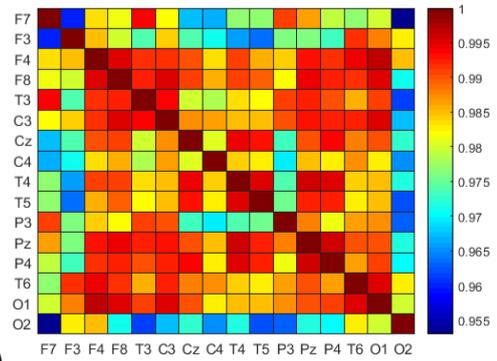
Table 3 Average L1

Channel	Control group (Average + SD)	Sch group (Average + SD)	P- value
'F7'	0,4213 ± 0,1196	0,4251 ± 0,1558	0,9173
'F3'	0,5704 ± 0,1467	0,5716 ± 0,1245	0,9722
'F4'	0,5683 ± 0,1556	0,5571 ± 0,1914	0,7964
'F8'	0,5554 ± 0,1540	0,5483 ± 0,1567	0,8395
'T3'	0,5064 ± 0,1887	0,5207 ± 0,2236	0,7638
'C3'	0,5557 ± 0,1501	0,5650 ± 0,1881	0,8122
'Cz'	0,5758 ± 0,1808	0,6015 ± 0,1436	0,5063
'C4'	0,5891 ± 0,1263	0,5781 ± 0,1811	0,7574
'T4'	0,5308 ± 0,1544	0,5098 ± 0,2154	0,6383
'T5'	0,5121 ± 0,1944	0,4800 ± 0,2267	0,5302
'P3'	0,5853 ± 0,1919	0,5511 ± 0,2116	0,4797
'Pz'	0,6177 ± 0,1463	0,6163 ± 0,1340	0,9665
'P4'	0,6021 ± 0,1981	0,6061 ± 0,1302	0,9145
'T6'	0,5504 ± 0,1525	0,5912 ± 0,0968	0,1977
'O1'	0,6908 ± 0,1045	0,5861 ± 0,1990	0,0106
'O2'	0,5971 ± 0,2503	0,6602 ± 0,1062	0,1958

Comparison of topographic images of EEG signal complexity obtained by different methods allows us to conclude that the subjects from the control group and the patients with schizophrenia syndromes have clear differences in activity readings in the cortical areas. Cross-correlation function was used to measure these differences. Cross-correlation was determined for the average LZC for all subjects for each of the

EEG channels. Fig. 3 shows the visualization of the cross-correlation function for both groups.

a)



b)

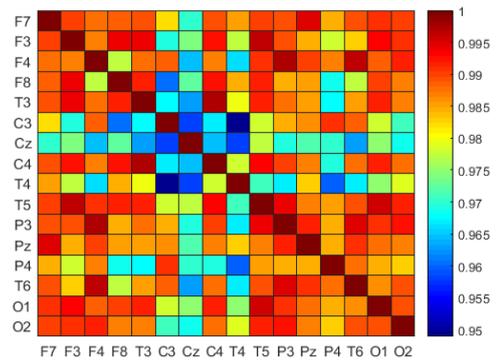


Fig. 3. Cross-correlation of LZC for a) control group and b) schizophrenia syndrome group

The following pairs of electrodes with the lowest value were identified: for the control group - O2-F8, O2-T3, O2-C4, O2-T4, O2-T5, O2-P3, O2-Pz, O2-P4; for the group with schizophrenia syndromes - C3-F8, C3-T3, Cz-T3, Cz-C3, C4-C3, C4-Cz, T4-C3, T4-Cz. Thus, a decrease in cross-correlation values in patients with schizophrenia was found for LZC in the left hemisphere compared to the right hemisphere, mainly in the parietal zone. The control group is characterized by weak cross-correlation, predominant right occipital zone with parietal zone. These results support the assumption that schizophrenia may be a disruption of the activity ratio in different brain regions.

#### 4. Conclusion

The proposed method of visual analysis of EEG allowed us to compare the interaction between the activity of brain regions. This approach makes it possible to evaluate the symmetry of activity on the basis of topographic images, to localize the activity centers and to correlate the activity of interaction between hemispheres by means of cross-correlation analysis.

The most characteristic EEG channels were selected for each method. Comparison of the methods for determining the signal complexity has shown that the most characteristic is LZC, because 5 significant channels were determined for this method.

The results will allow to implement the evaluation functionality with the use of machine learning for further research in medical diagnosis of schizophrenia.

#### 5. Acknowledgements

This research was supported by the Ministry of Science and Higher Education of the Russian Federation, project No. 3.861.2017 / 4.6.

## 6. References

1. Akar S. A. et al. Analysis of the complexity measures in the EEG of schizophrenia patients //International journal of neural systems. – 2016. – Т. 26. – №. 02. – С. 1650008.
2. Costa M., Goldberger A. L., Peng C. K. Multiscale entropy analysis of biological signals //Physical review E. – 2005. – Т. 71. – №. 2. – С. 021906.
3. Elvevag B., Goldberg T. E. Cognitive impairment in schizophrenia is the core of the disorder //Critical Reviews™ in Neurobiology. – 2000. – Т. 14. – №. 1.
4. Ferenets R. et al. Comparison of entropy and complexity measures for the assessment of depth of sedation //IEEE Transactions on Biomedical Engineering. – 2006. – Т. 53. – №. 6. – С. 1067-1077.
5. Gaebel W. et al. Trends in Schizophrenia Diagnosis and Treatment //Advances in Psychiatry. – Springer, Cham, 2019. – С. 603-619. Автор В. Источник.
6. Humeau-Heurtier A. The multiscale entropy algorithm and its variants: A review //Entropy. – 2015. – Т. 17. – №. 5. – С. 3110-3123.
7. Kaspar F., Schuster H. G. Easily calculable measure for the complexity of spatiotemporal patterns //Physical Review A. – 1987. – Т. 36. – №. 2. – С. 842.
8. Kim D. J. et al. An estimation of the first positive Lyapunov exponent of the EEG in patients with schizophrenia //Psychiatry Research: Neuroimaging. – 2000. – Т. 98. – №. 3. – С. 177-189.
9. Lempel A., Ziv J. On the complexity of finite sequences //IEEE Transactions on information theory. – 1976. – Т. 22. – №. 1. – С. 75-81.
10. Röscke J., Fell J., Beckmann P. Nonlinear analysis of sleep EEG data in schizophrenia: calculation of the principal Lyapunov exponent //Psychiatry research. – 1995. – Т. 56. – №. 3. – С. 257-269.
11. Wolf A. et al. Determining Lyapunov exponents from a time series //Physica D: Nonlinear Phenomena. – 1985. – Т. 16. – №. 3. – С. 285-317.
12. Wu X., Xu J. Complexity and brain function //Acta Biophysica Sinica. – 1991. – Т. 7. – №. 1. – С. 103-106.