Random matrix theory analysis of EEG data

Min Lei
Associate Professor, Institute of Vibration, Shock & Noise, State Key Laboratory of Mechanical System and Vibration, Fundamental Science on Vibration, Shock & Noise Laboratory, Department of Mechanical Engineering, Shanghai Jiao Tong University, China

Jingdong Li
Graduate Student, Institute of Vibration, Shock & Noise, State Key Laboratory of Mechanical System and Vibration, Fundamental Science on Vibration, Shock & Noise Laboratory, Department of Mechanical Engineering, Shanghai Jiao Tong University, China

ABSTRACT: In this study, we first construct the empirical cross-correlation matrices in the EEG signals by the Pearson's correlation coefficient. Then, we apply random matrix theory (RMT) investigate the statistical properties of cross-correlations in EEG data. The EEG signals of Alzheimer's disease (AD) subjects are chosen as the research objects. Here, we examine the statistical properties of cross-correlation coefficients, the distribution of eigenvalues. Meanwhile, we identify the deterministic or stochastic dynamics of the EEG system based on RMT. The correlation matrix of a Wishart matrix or a Gaussian Wigner matrix in random matrix theory has a structure similar to that of EEG data. For a noisy signal, its eigenvalue statistics are closely those of random matrix ensembles. The results show that the EEG signals are different from the stochastic time series.

1. INTRODUCTIONS

The electroencephalogram (EEG) is an important non-invasive technique to show the activity of the brain in real time. EEG data are often extremely complex and noisy although the EEG technique has been used in medical diagnosis and brain research. It is of crucial for understanding the brain state how to characterize EEG signals. For the aim of helping to explore the meaning of EEG signals, many methods have been developed from time domain to frequency domain, from linear analysis to nonlinear analysis, such as AR model, spectral powers in five wide frequency bands, entropy estimate, etc. However, it is still a challenge for these methods how to understanding EEG signals measured from abnormal subjects.

In statistics, random matrix theory (RMT) has successfully been proposed to explain the statistical properties of complex systems and multivariate time series, such as quantum chaotic systems, large complex atoms, price fluctuations in stock market, EEG data of brain, etc. For complex systems, their spectral properties have been predicted by RMT. For multivariate time series data, their empirical cross-correlation matrices have been analyzed in terms of RMT.

The purpose of this paper is to investigate the existence of properties of the EEG signals for Alzheimer's disease (AD). RMT has been applied to deal with the cross-correlations in EEG signals and compare them in different brain areas.

2. METHODS AND MATERIALS

2.1. Methodology

For eigenvalues spectra study of the random matrix, spectral density or distribution of eigenvalues have to be considered. Assuming that a time series x_1, x_2, Λ, x_n is coming from a Gaussian random system, it can be embedded into $d\times m$ dimension random matrix X in terms of the analysis period length:

$$X_{m\times d}=\begin{bmatrix} x_1 & x_{m+1} & \cdots & x_{(d-1)m+1} \\ x_2 & x_{m+2} & \cdots & x_{(d-1)m+2} \\ \vdots & \vdots & \cdots & \vdots \\ x_m & x_{2m} & \cdots & x_{dm} \end{bmatrix} \quad (1)$$

where d is dimension, m is the length of the period studied, $d{\cdot}m\leq n$. Thus, X can describe the characteristics of the system according to Takens' embedding theorem. The correlation matrix $C_{d\times d}(m)$ of the matrix X by Pearson's correlation coefficient C_{ij}:

$$C_{ij}=\frac{\langle X_i X_j\rangle-\langle X_i\rangle\langle X_j\rangle}{\sigma_i\sigma_j} \quad (2)$$

where <·> means the time average over the analysis period. X_i and X_j represent the ith and jth column vector in X, respectively. σ_i is the standard deviation of X_i. C_{ij} ranges between -1 and 1.The matrix $C_{d\times d}(m)$ describe the correlation between two different analysis period for a time series $\boldsymbol{x}$.

The eigenvalues λ_k (k=1,…,d) of the correlation matrix $C_{d\times d}$ is given by the singular value decomposition (SVD). The eigenvalue distribution $\rho(\lambda)$ is the simplest property of the eigenvalues family. In terms of the random matrix theory, universal statistical properties of its correlation matrix C can be predicted by the eigenvalues spectra. That is, for random real symmetric matrices (or Gaussian orthogonal ensemble, GOE), the eigenvalues probability distribution is well approximated by the Wigner formal:

$$\rho_{\mathrm{GOE}}(\lambda)=\frac{\pi}{2}\lambda\exp\left(-\frac{\pi}{4}\lambda^2\right) \quad (3)$$

Here, the distribution is a Gaussian probability distribution.

2.2. *EEG data*

The analyzed EEG signals are collected during the resting state with open eyes for the Alzheimer's disease (AD), mild cognitive impairment (MCI) and healthy subjects, respectively. The sampling frequency is 1kHz. There are 16 channel electrodes located at the scalp positions of Fp1, Fp2, F3, F4, C3, C4, P3, P4, F7, F8, T3, T4, T5, T6, O1, O2 (see Fig.1). Here, the EEG signal of each channel is denoted as x_i, i = 1, … , 50000. Because the electrodes are very close, the correlation between EEG signals of each channel is strong. That means that we cannot get a like random matrix from all channel data. Here, we analyze the EEG correlations of single channel to investigate the dynamical characteristics of the brain activity. The matrix X based on EEG data is constructed at m=2000, d=25.

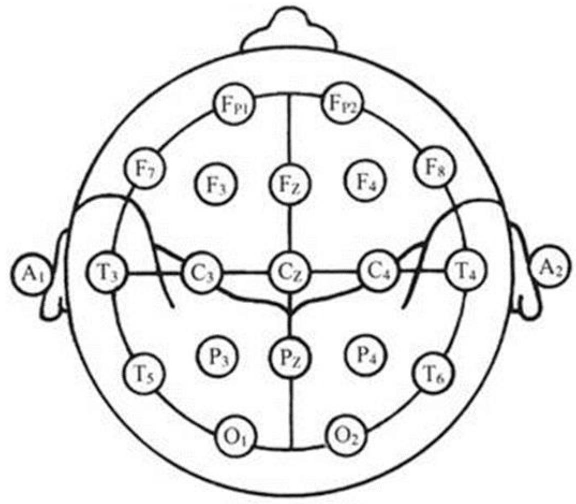
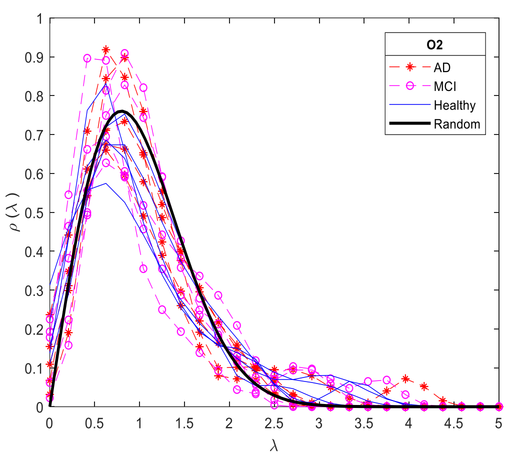


Figure 1: 10−20 system.

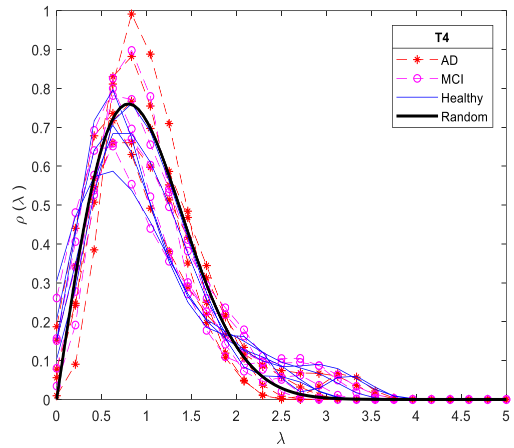
3. ANALYSIS AND RESULTS

The eigenvalues distribution of EEG signal is dependent on the underlying brain activity system because EEG synchronizes with the corresponding brain activity. We assume that the brain activity is random during the resting state and open eyes. Thus, the empirical cross-correlation matrix C is considered consistent with a real-symmetric random matrix. We investigate the EEG signals of three types of subjects, five in each category and evaluate the density function of the corresponding channels.

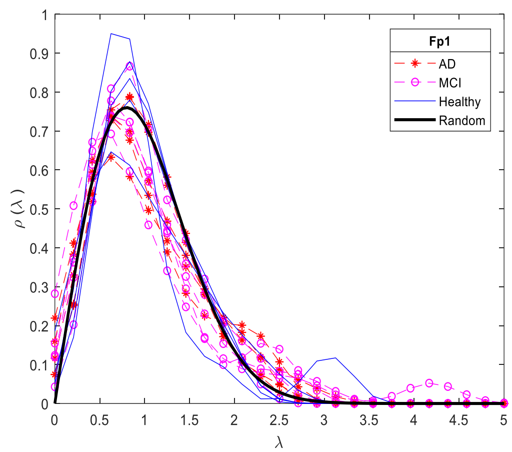
Figure 2 shows the eigenvalue distributions $\rho(\lambda)$ of correlation matrix C of EEG data for AD, MCI and healthy subjects at O2 (see Fig. 2a), T4 (see Fig. 2b), Fp1(see Fig. 2c) and Fp2(see Fig.2d) positions, respectively, as well as the eigenvalue distribution $\rho_{\mathrm{GOE}}(\lambda)$ of a random real symmetric matrices. Compared with the eigenvalue distributions of healthy subjects, we can see that some eigenvalue distributions of correlation matrix C of EEG signals for AD and MCI subjects are not close to the eigenvalue distribution of the random matrix when the eigenvalues are around one. For AD and MCI subjects, their eigenvalue distributions at O2 and T4 positions are either higher than random distribution values or lower than random values (see Fig.2a and 2b). These results might imply that for AD and MCI subjects, the brain function area at O2 and T4 are different from those of healthy subjects. At Fp1 and Fp2 positions, the eigenvalue distributions of correlation matrix *C* of EEG data for healthy subjects are similar to those for AD and MCI subjects (see Fig.2c and 2d). And the results are almost all around the eigenvalue distribution of the random matrix. It indicates that the correlation matrix of EEG data at Fp1 and Fp2 positions are roughly consistent with the random matrix distribution. At Fp1 and Fp2, the EEG data could be uncorrelated time series. Comparing Fig. 2a and 2b with Fig. 2c and 2d, the brain functions of occipital and temporal lobes might be abnormal for AD and MCI subjects.



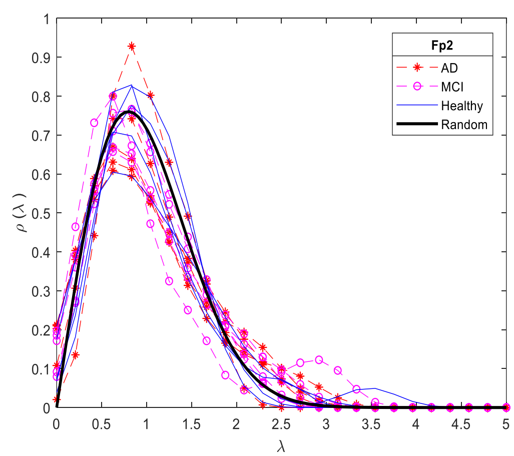
(*a*)



(*b*)



(*c*)



(*d*)

Figure 2: The eigenvalues distribution P(λ) of correlation matrix C of EEG data for three types of subjects at (a) O2, (b) T4, (c) Fp1, (d) Fp2 position (Red "" for AD subjects, magenta "o" for MCI subjects, blue solid line for healthy subject, black solid line for the Wigner distribution by Eq.（3）).*

4. CONCLUSIONS

The results presented above are agree well with RMT predictions for the EEG signals at Fp1 and Fp2 positions. But, for AD and MCI subjects, the eigenvalue distributions of the EEG signals at O2 and T4 positions deviate from RMT. The results show that the EEG signals are different from the stochastic time series. These indicate that the systems are specific and contain collective modes. AD and MCI subjects maybe have brain dysfunction in occipital and temporal lobes.

5. REFERENCES

Jalan, S., Bandyopadhyay, JN.(2007) "Random matrix analysis of complex networks" *Physical Review E*, 76, 046107

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L.A.N., Stanley, H.E(2000). "A random matrix theory approach to financial cross-correlations" Physica A, 287, 374-382.

Sakkalis, V., Giurcăneanu, C., Xanthopoulos, P., Zervakis, M., Tsiaras, V., Yang, Y., Karakonstantaki, E., Micheloyannis, S.(2009). "Assessment of linear and nonlinear synchronization measures for analyzing EEG in a mild epileptic paradigm" IEEE transactions on information technology in biomedicine, 13(4), 433-441.

Šeba, P. (2003). "Random matrix analysis human EEG data" *Physical Review Letters*,91(19), 198104-1–4.

Wang, G., Xie, C., Chen, S., Yang, J., Yang, M.(2013) "Random matrix theory analysis of cross-correlations in the US stock market: evidence from Pearson's correlation coefficient and detrended cross-correlation coefficient" Physica A, 392, 3715-3730.

Ye, B., Chen, J., Ju, C., Li, H., Wang, X. (2017). "Distinguishing chaotic time series from noise: a random matrix approach" Commun Nonlinear Sci Numer Simulat, 44, 284-291.