This is the Author's *accepted* manuscript version of the following contribution:

[Matteo Demuru and Matteo Fraschini, **EEG fingerprinting: subject-specific signature based on the aperiodic component of power spectrum, Computers in Biology and Medicine, Volume 120, May 2020, 103748**]

The publisher's version is available at:
https://doi.org/10.1016/j.compbiomed.2020.103748

When citing, please refer to the published version.

EEG fingerprinting: subject-specific signature based on the aperiodic component of power spectrum

Matteo Demuru[1] and Matteo Fraschini[2]

[1] Stichting Epilepsie Instellingen Nederland (SEIN), Heemstede, The Netherlands
[2] Department of Electrical and Electronic Engineering, University of Cagliari, Cagliari, Italy

Corresponding author: Matteo Fraschini (e-mail: fraschin@unica.it).

Abstract

During the last few years, there has been growing interest in the effects induced by individual variability on activation patterns and brain connectivity. The practical implications of individual variability are of basic relevance for both group level and subject level studies. The Electroencephalogram (EEG), still represents one of the most used recording techniques to investigate a wide range of brain-related features. In this work, we aim to estimate the effect of individual variability on a set of very simple and easily interpretable features extracted from the EEG power spectra. In particular, in an identification scenario, we investigated how the aperiodic (1/f background) component of the EEG power spectra can accurately identify subjects from a large EEG dataset. The results of this study show that the aperiodic component of the EEG signal is characterized by strong subject-specific properties, that this feature is consistent across different experimental conditions (eyes-open and eyes-closed) and outperforms the canonically-defined frequency bands. These findings suggest that the simple features (slope and offset) extracted from the aperiodic component of the EEG signal are sensitive to individual traits and may help to characterize and make inferences at single subject-level.

Keywords: EEG; fingerprint; power spectra; aperiodic component.

Introduction

The scalp electroencephalogram (EEG) represents the most common recording system to detect human brain activity. Despite that this technique dates back to more than a century ago, it remains a clear reference in all the neuroscience related fields (Lopes da Silva 2013). One of the most relevant advantages of EEG, compared to other well-known techniques as Magnetoencephalography (MEG) and functional magnetic resonance (fMRI), is represented by its relatively low-cost and its non-invasive properties that make it especially suitable for real-life applications. Thanks to these characteristics, more recently, scalp EEG has been extensively investigated as source of neurophysiological features to be used in the so-called cognitive biometric systems (Chan et al. 2018; Gui et al. 2019). To date, an impressive number of feature extraction techniques have been used to detect subject-specific traits from EEG recordings. Time, frequency, time-frequency -domain techniques, functional connectivity metrics, and the related derived network parameters have been extensively investigated into this specific research topic. Among these techniques, the most frequently used derive either from the use of power spectral density analysis at channels level (DelPozo-Banos et al. 2015; Rocca et al. 2014) or from functional connectivity between EEG channels (Barra et al. 2017; Crobe et al. 2016; Fraschini et al. 2015). All these approaches, from the most simple to the most sophisticated, are based on some sort of arbitrary choice, such as (i) the frequency band definition, (ii) the selection of a correlation metric for the connectivity and/or (iii) the application of a threshold to reconstruct the functional network. Interestingly, it has been shown that EEG brain activity exhibits a 1/f-like power spectrum (He 2014), defining an aperiodic component that may be characterized in terms of slope (i.e., the exponential decrease of power in a spectrogram as a function of frequency) and offset (i.e., offset of the broadband power of the signal). This arrhythmic brain activity has been associated with fluctuations of cognitive states (Podvalny et al. 2015), aging (Voytek et al. 2015), the firing rate of neural populations (Manning et al. 2009) and clinical conditions (Peterson et al. 2018; Robertson et al. 2019; Veerakumar et al. 2019). Despite the interest in this specific approach, to the best of our knowledge, it has not been investigated yet to what extent this aperiodic component depicts the individual variability in EEG brain activity. In this study, we quantified the aperiodic component of the power spectrum, estimating the spectral slope and the offset at channel level. Successively, we used these EEG characteristics as feature vector to identify subjects in a large EEG dataset. Finally, we compared the results obtained from the aperiodic component with those obtained from more classical power spectral features, namely the relative power extracted from the canonical EEG frequency bands (i.e., theta, alpha, beta and gamma bands). In order to investigate this question, we used resting-state traces from a large and publicly available EEG dataset consisting of several recordings from 109 different subjects using a 64 channel EEG equipment (Goldberger Ary L. et al. 2000; Schalk et al. 2004).

Material and methods

Dataset

To test our hypothesis we use a large and publicly available (https://physionet.org/content/eegmmidb/1.0.0/) EEG dataset consisting of several recordings from 109 different subjects using high-density EEG equipment (Goldberger Ary L. et al. 2000; Schalk et al. 2004). The EEG traces are provided in EDF+ format, containing 64 EEG signals and sampled at 160 samples per second. In particular, since the possible effects on individual variability of different sessions and different tasks have already been investigated (Fraschini et al. 2019; Pani et al. 2019) and are out of the scope of the present work, we focused our analysis on the two (one minute long) resting-state recordings during eyes-closed and eyes-open conditions.

Pre-processing

The pre-processing procedure was organized into two main steps. The first step was to apply ADJUST (version 1.1.1, https://www.nitrc.org/projects/adjust/), an automatic algorithm for artifact identification and removal (Mognon et al. 2011) based on Independent Component Analysis (ICA), with the aim to reduce the effects due to blinks, eye movements, and other generic discontinuities. The subjects that showed significant residual artifacts were excluded from the analysis; in particular, the reported results are based on 95 subjects for the eyes-open condition and 100 subjects for the eyes-closed condition. The second step was to segment each (one minute) resting-state EEG recordings into five non-overlapping epochs of 12 seconds (Fraschini et al. 2015; Fraschini et al. 2016).
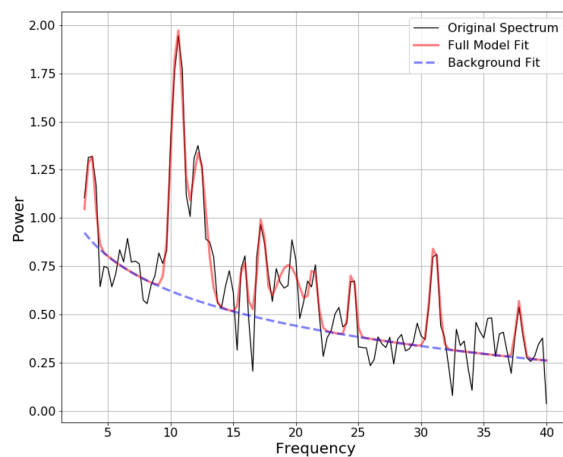

Figure 1. The original power spectrum (extracted from one subject – one channel, using real EEG data) with the associated FOOOF model fit which allowed to compute both the slope and the offset for the aperiodic component.

Features extraction

After the pre-processing steps, two different types of features were extracted from the epoched signals: (i) features characterizing the aperiodic component, namely the slope and the offset and (ii) features characterizing the periodic component, namely the relative power of theta [4-8 Hz], alpha [8-13 Hz], beta [13-30 Hz] and gamma [30-45 Hz] frequency bands. The slope and the offset were calculated using the Fitting Oscillations & One Over f (FOOOF) toolbox (Haller et al. 2018). The slope and offset are derived modelling the aperiodic

signal (*L*), see Figure 1, as $L = b - log(k + F^x)$, where b is the offset, x the slope and k the so-called knee parameter (that controls the bend of the aperiodic component). F represents a 1-D vector containing the frequencies. The relative power for each frequency band has been computed as the ratio between the absolute band-specific power and absolute total power (between 1 and 45 Hz) using the Power Spectral Density estimate via Welch's method. Both approaches allowed to obtain, for each subject and each epoch, a features vector of 64 entries (each one representing the corresponding slope, offset or relative power of a single EEG channel).

Performance evaluation

To assess the performance of the two approaches (periodic and aperiodic components of the power spectra) to capture subject-specific characteristics, we have tested the extracted features using a well-known paradigm generally used to evaluate biometric systems (Fraschini et al. 2015; Ross et al. 2006). In particular, genuine and impostor scores were computed based on the Euclidean distance between pairs of feature vectors. Later, a similarity score was computed as 1/(1+d), where d represents the Euclidean distance. Finally, the performance is derived from the false acceptance rate (FAR, the error that occurs when the system accepts an impostor) and the false rejection rate (FRR, the error that occurs when the system rejects a genuine match) at different thresholds. The area under the receiver operating characteristic (AUC) curves were evaluated together with the equal error rate (EER, the point where FAR equals FRR) and the correct recognition rate (CRR) and were reported to summarize the results. Low values of EER and AUC express high performance. All the analysis was performed using MATLAB (The MathWorks, Inc., Natick, Massachusetts, United States, version 2017B) and EEGLAB (version 13) (Delorme and Makeig 2004). All the figures were realized using Jamovi (version 1.0.8.0) available from https://www.jamovi.org.

Results

Eyes-open resting-state

For the eyes-open resting-state condition, the best performance, in terms of EER and AUC, were obtained for the offset (EER = 0.079 and AUC = 0.025). The slope performed slightly worse, with EER = 0.090 and AUC = 0.031. Figure 2 represents the similarity score distribution for the slope (left panel) and the offset (right panel). When the two feature vectors were concatenated, the performance was slightly better with EER = 0.063 and AUC = 0.019. The relative power feature vectors did not show comparable performance, with the best results obtained for beta (EER = 0.118 and AUC = 0.043) and gamma band (EER = 0.112 and AUC = 0.039). All the results for the eyes-open resting-state condition are summarized in Table 1.
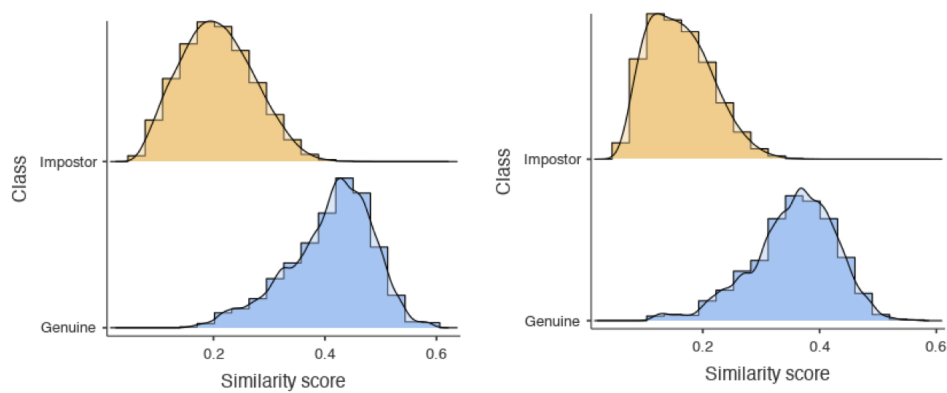

Figure 2. Similarity score distribution for the slope (left panel) and the offset (right panel) for eyes-open resting-state condition

Feature	EER	CCR
slope	0.090	0.962
offset	0.079	0.960
slope + offset	0.063	0.973
theta relative power	0.182	0.719
alpha relative power	0.256	0.762
beta relative power	0.118	0.938
gamma relative power	0.112	0.942
all bands	0.101	0.966

Table 1. A summary of the EER values for each feature vector for eyes-open resting-state condition.

Eyes-closed resting-state

For the eyes-closed resting-state condition, the best performance, in terms of EER and AUC, were obtained for the offset (EER = 0.057 and AUC = 0.018). The slope performed slightly worse, with EER = 0.089 and AUC = 0.035. Figure 3 represents the similarity score distribution for the slope (left panel) and the offset (right panel). In this case, when the two feature vectors were concatenated, the performance was slightly worse with EER = 0.171 and AUC = 0.019. Again, the relative power feature vectors did not show comparable performance, with the best results obtained for beta (EER = 0.142 and AUC = 0.078) and gamma band (EER = 0.197 and AUC = 0.124). All the results for the eyes-closed resting-state condition are summarized in Table 2. Finally, the distributions of slope and offset for two sample subjects (S002R02 and S003R02) during the eyes-closed resting state for all the 64 EEG channels are represented in Figure 4.
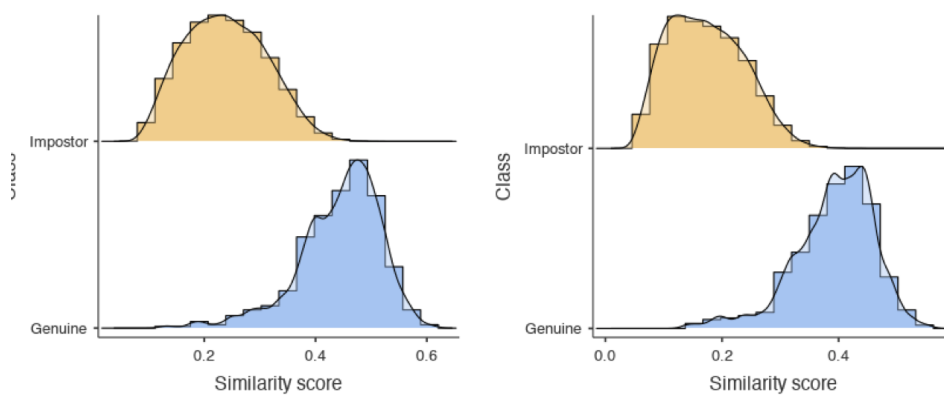

Figure 3. Similarity score distribution for the slope (left panel) and the offset (right panel) for eyes-closed resting-state condition

Feature	EER	CRR
slope	0.089	0.940
offset	0.057	0.960
slope + offset	0.171	0.876
theta relative power	0.238	0.630
alpha relative power	0.181	0.813
beta relative power	0.142	0.877
gamma relative power	0.197	0.947
all bands	0.119	0.940

Table 2. A summary of the EER values for each feature vector for eyes-closed resting-state condition.
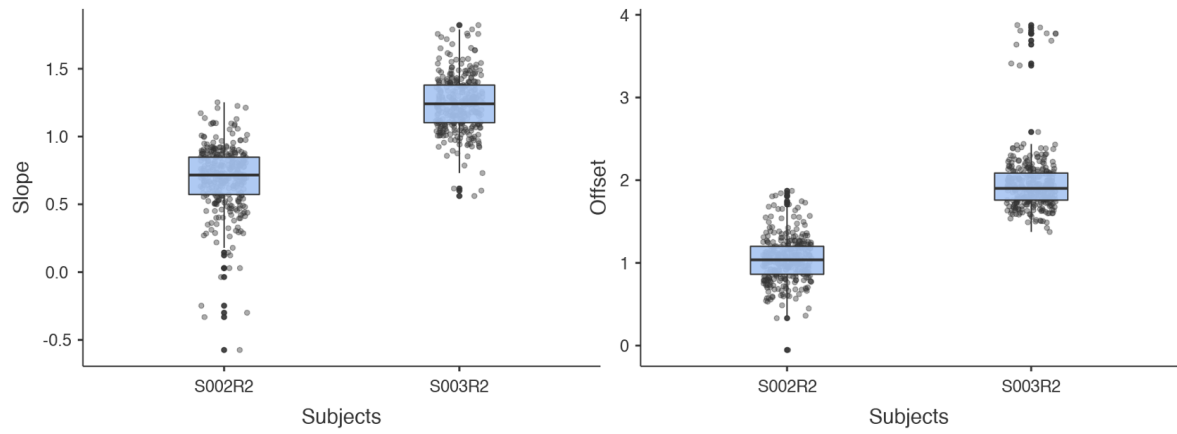
Figure 4. Left and right panels represent respectively the distributions of slope and offset for two sample subjects (S002R02 and S003R02) during eyes-closed resting state for all the 64 EEG channels.

Discussion

In this study, a large and open EEG dataset is used to investigate inter-subject variability using a fast and easy approach that goes beyond the traditional frequency band analysis. Two simple characteristics (slope and offset of the broadband power spectrum) describing the aperiodic properties of the EEG signals were exploited as fingerprints to identify individuals. Our results show that the aperiodic component of the power spectrum (i) is characterized by strong subject-specific properties, (ii) is consistent across different experimental conditions (eyes-open and eyes-closed) and (iii) its identification performance outperforms those obtained using canonical spectral features as band-specific relative power.

Overall, the best performance is obtained using the aperiodic offset in the eyes-closed resting-state condition. Nevertheless, there is not a marked difference in terms of performance between experimental conditions (eyes-open versus eyes-closed resting-state) for any of the feature vectors tested. In line with previous studies (Crobe et al. 2016; Fraschini et al. 2015), for the periodic component, we observed an effect on the performance due to the high-frequency content, where for both conditions the lowest EER values (best performance) were obtained for beta and gamma frequency bands. It is interesting to notice that when combining the features derived from all the frequency bands, the results outperform those obtained considering the bands separately, but still, both the slope and offset remain the features showing the better performance. The fact that the aperiodic component of the power spectrum outperforms the traditional spectral features (i.e., band-specific relative power) is in line with the hypothesis that the a priori and arbitrary frequency bands definition and the averaging across these bands can mitigate individual variability (Haller et al. 2018). Furthermore, these results also support the hypothesis that this approach may help to better understand the role of oscillatory variability in explaining individual differences in cognitive functioning in health and disease (Haller et al. 2018).

Moreover, the aperiodic component has been linked experimentally to neuronal processes at micro-scale level (Manning et al. 2009; Miller et al. 2012; Podvalny et al. 2015) and it is hypothesized that its physiological meaning might reflect the dynamic balance between excitation and inhibition of neural population (Gao et al. 2017). In this work, we tried to make a bridge testing if properties reported at micro/mesoscale were reliable at the macroscale (i.e. EEG recordings). Similarly to our results, recent works at the macroscale level showed the reliability of individual functional brain connectivity profile (Finn et al. 2015) and its strong genetic dependence (Demuru et al. 2017). A clear limitation of the present study is that it is based on a single session scenario and it is not clear how these findings may be altered using different EEG recordings in a multi-session scenario. Nevertheless, it is relevant to notice that our work is especially focused on relative (and not absolute) performance since we present the results as a strict comparison between different characteristics of the EEG spectrum (namely, aperiodic and periodic components). Furthermore, the effects induced by different sessions on subject identification in EEG have been widely investigated and clarified (Gui et al. 2019) and are out of the scope of the present study. An important limitation of the present work is that we have not considered the possible effects induced by the reference used in the analysis. We think this point may deserve more attention in future studies.

Conclusions

In conclusion, in this study, we have shown that an EEG individual's profile, as defined by the aperiodic component of the power spectrum, is unique and it is possible to identify individuals (with very high accuracy) from a large EEG dataset. These findings suggest that these simple spectral features are sensitive to individual traits and may help to characterize and make inferences at single subject-level using EEG.

Authors' contributions

Matteo Demuru and Matteo Fraschini conceived and designed the study, analyzed and contributed equally to the writing the final manuscript.

Declaration of competing interest

The authors have no conflict of interest to disclose.

References

Barra, S., Casanova, A., Fraschini, M., & Nappi, M. (2017). Fusion of physiological measures for multimodal biometric systems. *Multimedia Tools and Applications*, *76*(4), 4835–4847. https://doi.org/10.1007/s11042-016-3796-1

Chan, H.-L., Kuo, P.-C., Cheng, C.-Y., & Chen, Y.-S. (2018). Challenges and Future Perspectives on Electroencephalogram-Based Biometrics in Person Recognition. *Frontiers in Neuroinformatics*, *12*. https://doi.org/10.3389/fninf.2018.00066

Crobe, A., Demuru, M., Didaci, L., Marcialis, G. L., & Fraschini, M. (2016). Minimum spanning tree andk-core decomposition as measure of subject-specific EEG traits. *Biomedical Physics & Engineering Express*, *2*(1), 017001. https://doi.org/10.1088/2057-1976/2/1/017001

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

DelPozo-Banos, M., Travieso, C. M., Weidemann, C. T., & Alonso, J. B. (2015). EEG biometric identification: a thorough exploration of the time-frequency domain. *Journal of Neural Engineering*, *12*(5), 056019. https://doi.org/10.1088/1741-2560/12/5/056019

Demuru, M., Gouw, A. A., Hillebrand, A., Stam, C. J., Dijk, B. W. van, Scheltens, P., et al. (2017). Functional and effective whole brain connectivity using magnetoencephalography to identify monozygotic twin pairs. *Scientific Reports*, *7*(1), 9685. https://doi.org/10.1038/s41598-017-10235-y

Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., et al. (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature Neuroscience*, *18*(11), 1664–1671. https://doi.org/10.1038/nn.4135

Fraschini, M., Hillebrand, A., Demuru, M., Didaci, L., & Marcialis, G. L. (2015). An EEG-Based Biometric System Using Eigenvector Centrality in Resting State Brain Networks. *IEEE Signal Processing Letters*, *22*(6), 666–670. https://doi.org/10.1109/LSP.2014.2367091

Fraschini, Matteo, Demuru, M., Crobe, A., Marrosu, F., Stam, C. J., & Hillebrand, A. (2016). The effect of epoch length on estimated EEG functional connectivity and brain network organisation. *Journal of Neural Engineering*, *13*(3), 036015. https://doi.org/10.1088/1741-2560/13/3/036015

Fraschini, Matteo, Pani, S. M., Didaci, L., & Marcialis, G. L. (2019). Robustness of functional connectivity metrics for EEG-based personal identification over task-induced intra-class and inter-class variations.

Pattern Recognition Letters, *125*, 49–54. https://doi.org/10.1016/j.patrec.2019.03.025

Gao, R., Peterson, E. J., & Voytek, B. (2017). Inferring synaptic excitation/inhibition balance from field potentials. *NeuroImage*, *158*, 70–78. https://doi.org/10.1016/j.neuroimage.2017.06.078

Goldberger Ary L., Amaral Luis A. N., Glass Leon, Hausdorff Jeffrey M., Ivanov Plamen Ch., Mark Roger G., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet. *Circulation*, *101*(23), e215–e220. https://doi.org/10.1161/01.CIR.101.23.e215

Gui, Q., Ruiz-Blondet, M. V., Laszlo, S., & Jin, Z. (2019). A Survey on Brain Biometrics. *ACM Computing Surveys (CSUR)*, *51*(6), 112:1–112:38. https://doi.org/10.1145/3230632

Haller, M., Donoghue, T., Peterson, E., Varma, P., Sebastian, P., Gao, R., et al. (2018). Parameterizing neural power spectra. *bioRxiv*, 299859. https://doi.org/10.1101/299859

He, B. J. (2014). Scale-free brain activity: past, present, and future. *Trends in Cognitive Sciences*, *18*(9), 480–487. https://doi.org/10.1016/j.tics.2014.04.003

Lopes da Silva, F. (2013). EEG and MEG: relevance to neuroscience. *Neuron*, *80*(5), 1112–1128. https://doi.org/10.1016/j.neuron.2013.10.017

Manning, J. R., Jacobs, J., Fried, I., & Kahana, M. J. (2009). Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(43), 13613–13620. https://doi.org/10.1523/JNEUROSCI.2041-09.2009

Miller, K. J., Hermes, D., Honey, C. J., Hebb, A. O., Ramsey, N. F., Knight, R. T., et al. (2012). Human Motor Cortical Activity Is Selectively Phase-Entrained on Underlying Rhythms. *PLOS Computational Biology*, *8*(9), e1002655. https://doi.org/10.1371/journal.pcbi.1002655

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, *48*(2), 229–240. https://doi.org/10.1111/j.1469-8986.2010.01061.x

Pani, S. M., Ciuffi, M., Demuru, M., Bazzano, G., D'aloja, E., & Fraschini, M. (2019). Subject, session and task effects on power, connectivity and network centrality: a source-based EEG study. *bioRxiv*, 673343. https://doi.org/10.1101/673343

Peterson, E. J., Rosen, B. Q., Campbell, A. M., Belger, A., & Voytek, B. (2018). 1/f neural noise is a better predictor of schizophrenia than neural oscillations. *bioRxiv*, 113449. https://doi.org/10.1101/113449

Podvalny, E., Noy, N., Harel, M., Bickel, S., Chechik, G., Schroeder, C. E., et al. (2015). A unifying principle

underlying the extracellular field potential spectral responses in the human cortex. *Journal of Neurophysiology*, *114*(1), 505–519. https://doi.org/10.1152/jn.00943.2014

Robertson, M. M., Furlong, S., Voytek, B., Donoghue, T., Boettiger, C. A., & Sheridan, M. A. (2019). EEG power spectral slope differs by ADHD status and stimulant medication exposure in early childhood. *Journal of Neurophysiology*, *122*(6), 2427–2437. https://doi.org/10.1152/jn.00388.2019

Rocca, D. L., Campisi, P., Vegso, B., Cserti, P., Kozmann, G., Babiloni, F., & Fallani, F. D. V. (2014). Human Brain Distinctiveness Based on EEG Spectral Coherence Connectivity. *IEEE Transactions on Biomedical Engineering*, *61*(9), 2406–2412. https://doi.org/10.1109/TBME.2014.2317881

Ross, A. A., Nandakumar, K., & Jain, A. K. (2006). *Handbook of Multibiometrics*. Springer US. https://doi.org/10.1007/0-387-33123-9

Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering*, *51*(6), 1034–1043. https://doi.org/10.1109/TBME.2004.827072

Veerakumar, A., Tiruvadi, V., Howell, B., Waters, A. C., Crowell, A. L., Voytek, B., et al. (2019). Field potential 1/f activity in the subcallosal cingulate region as a candidate signal for monitoring deep brain stimulation for treatment-resistant depression. *Journal of Neurophysiology*, *122*(3), 1023–1035. https://doi.org/10.1152/jn.00875.2018

Voytek, B., Kramer, M. A., Case, J., Lepage, K. Q., Tempesta, Z. R., Knight, R. T., & Gazzaley, A. (2015). Age-Related Changes in 1/f Neural Electrophysiological Noise. *Journal of Neuroscience*, *35*(38), 13257–13265. https://doi.org/10.1523/JNEUROSCI.2332-14.2015