Proceedings of the 10th World Congress on Mechanical, Chemical, and Material Engineering (MCM'24)
Barcelona, Spain - August 22-24, 2024
Paper No. ICMIE 153
DOI: 10.11159/icmie24.153

Topological Data Analysis and Decoding Based on Electroencephalogram Signals

Shijin Xu[1], Jinyuan Song[1], Yiming Zhang[1], Jingyu Yang[1], Zheng Yang[1], Fok Sai Cheong[1], Peng Chen[2]
[1]Sichuan University-Pittsburgh Institute,
Sichuan University Jiang'an Campus, Chengdu, China
shijnxu26@gmail.com; 1097537934@qq.com; 1294120660@qq.com; 1733942830@qq.com;
zhengyang2018@scu.edu.cn; saicheong.fok@scupi.cn
[2]Southwest Jiaotong University, School of Mechanical Engineering
Southwest Jiaotong University Xipu Campus, Chengdu, China
chenpeng@swjtu.edu.cn

Abstract – In order to assist patients in performing upper limb activities in a normal manner, research on bimanual coordination is necessary for robots to assist in the rehabilitation process. The objective of this paper is to investigate brain activities in bimanual movements coordination to obtain and decode signals of the electroencephalogram (EEG) in bimanual movements in both real and virtual environments. We designed three paradigms of bimanual coordination tasks to collect and analyse EEG signals, including (1) raise both hands vertically, (2) spread both hands horizontally, (3) left hand horizontally spread to the left, right hand vertically rise upwards. We aim to classify the three types of the motions using persistent homology-based machine learning techniques. First we extract topological features from functional connectivity network and effective connectivity network based on correlations between channels. In particular, the persistent diagrams and persistent landscapes are calculated. Then we use them as input for the classification. Our results show that the highest average binary classification accuracies which is functional connectivity network (FC)+Persistent Scale Space kernel (PSSK), has an accuracy of only 72.48%±1.07% in the real environment. The design of our experimental paradigm and the classification results demonstrated the preliminary exploration and indicated the feasibility of decoding the bimanual coordination on EEG. In the future, we would expand the size of subjects and improve the topological data analysis methods for further exploration of EEG in the field of decoding two-handed coordination.

***Keywords*:** Electroencephalogram, Bimanual Coordination, Topological Data Analysis, Brain Networks, SVM

1. Introduction

In everyday life, two hands are usually used together to perform certain tasks or work. Therefore, bimanual movements are crucial and quite common in everyday activities. Compared to single-handed movements, bimanual coordination is gradually attracting more attention from researchers in the field of robotic rehabilitation [1]. How to record, analyse and decode the motion parameters of bimanual movements has become an important issue, which has not been thoroughly studied. Despite the success of the techniques such as invasive BCI in decoding bimanual coordination movements concerning brain actions, there are also some limitations to the security of such systems. Furthermore, the potential limitations of invasive BCI in the long term may restrict their clinical applications [2].

Electroencephalogram (EEG) signals and electromyography (EMG) signals are the most used methods to decode movements. Compared to using EMG signals (electromyogram, the recorded bioelectricity pattern of muscles), using EEG signals could detect movement intention faster. Wang et al. summarise that decoding hand movement parameters from EEG signals is feasible. We can extract hand kinematic information from EEG signals and decode movement parameters such as direction, acceleration, velocity and trajectory. Therefore, EEG-based movement intention decoding is an important branch in brain-computer interfaces (BCIs) [3].

After collecting EEG data, the next step is to decode the EEG signals into useful information. Topological Data Analysis (TDA) is one of the approaches that can extract EEG features. TDA is very suitable for the analysis of brain connectivity networks in neuroscience, and the key to the analysis is the topological structure of the brain network [4]. The TDA method will be used to analyse the EEG network as a method to decode EEG signals.

In a nutshell, bimanual coordination is a fundamental aspect of daily life. This study proposes an experimental paradigm based on EEG that aims to simplify and decode everyday bimanual movements. The main contributions of this study include: (1) design the motor execution experimental paradigms to collect and analyze EEG data during bimanual coordination tasks; (2) After preprocessing by using independent component analysis (ICA), re-reference, filtering and so on, use brain networks to analyze the strength of the EEG channels' relationships and proceeds to apply the persistent homology method chosen to obtain the persistent diagrams before finally inputting them into the kernel function based SVM machine learning method to obtain the accuracy.

The paper is organized as follows: Section 2 outlines the methodology employed in this paper. It describes the experimental paradigm, the human participants, and the data processing, with the data analysis by TDA and SVM. Section 3 details the results of brain network analysis and classification, which present in the average accuracy of the three subjects with the combination of the methods. Section 4 shows the discussion and then summarizes the results of our paper and suggests some areas of improvement based on the results.

2. Methodology

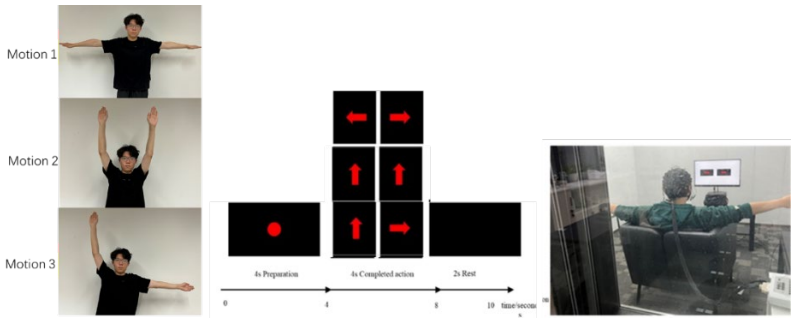
2.1. Experimental Paradigm and Experimental Execution

Because most bimanual movements can be simplified as the coordination movements of the two hands in these two directions. For the design of the experimental process, we aimed to make experiments effective and concise. We divided the common bimanual movements into three kinds and simplify them, and designed experiments based on these three types of movements. The design included three types of bimanual coordination: horizontal movement (M1); vertical movement (M2); horizontal + vertical movement (M3). The three movements are shown in the diagram below Figure 1 (a).

In a fully designed experiment, three healthy subjects (3 males), aged between 20-22, participated in the experiment. They were all right-handed and had normal or corrected-to-normal vision. The number and sequence of movements in each group of experiments were designed as a total of 30 movements in one group of experiments, each movement being performed 10 times, and each subject completing 5 groups of movements, with a 2-minute break in the middle, to avoid fatigue caused by repetition of the experiments.

During the experiment, the tasks informed participants of their actions on the screen. In the preparation phase, there was a be red cross on the screen to arouse the attention, and the process lasted 4s. After then, two red eye-catching arrows appeared on the screen, which located on the left and right sides, corresponding to the actions of the left and right hands. The arrows point in four directions (front, back, left and right), and the subjects completed the corresponding actions once according to the prompts. The process lasted for 4s. In the rest stage, the screen prompted the subjects to take a short rest, and the subjects could move their hands back to the original position and waited for the next action, which lasted for 2s. See Fig. 1 (a) and (b). And figure of the experiment in progress is shown in Fig. 1 (c).The experiment was conducted in an electromagnetically shielded room. Subjects were asked to sit on a chair in front of a table. A monitor was placed on the table, which randomly displayed the requested rectangular target action and was placed at a comfortable viewing distance of about 2 meters from the subject.

Prior to the formal experiment, subjects received approximately 15 minutes of pre-training on the task and were asked to avoid blinking and body movements during the experiment. The formal experiment was organized as described previously. Subjects were strictly advised to avoid any physical exercise while the EEG was being recorded.



(a) (b) (c)

Fig.1: Experimental setup (a) Illustration of three movements(b) Experimental process of bimanual coordination (on-screen screen guidance display) (c) Experiment in progress returned figure

2.2. Data Acquisition

The experiment was conducted in the laboratory provided by the School of Literature and Journalism in Sichuan University. The EEG cap has 66 channels, with sampling rate 500 Hz, which means it will collect 500 data points in 1s. There are ECOG electrode on the lower eyelid and the reference electrode on the tip of the nose and reference channels. Those two electrodes were processed during re-reference. In accordance with the International 10-10 System, the electrode locations were designated as such, with a forehead ground electrode at AFz. The data collected in the experiment was saved in .cdt format. The data preprocessing was done in real and virtual reality environment. We refer to [5] for details.

2.3. Data Analysis

The input data after preprocessing of one movement was a matrix of size 64*2500, where 64 represents the number of channels after re-reference and 2500 represents the total number of data points (500Hz sampling rate and 5 seconds, 2500 = 5*500). For the data analysis, the overall methodology was presented in Fig.2.

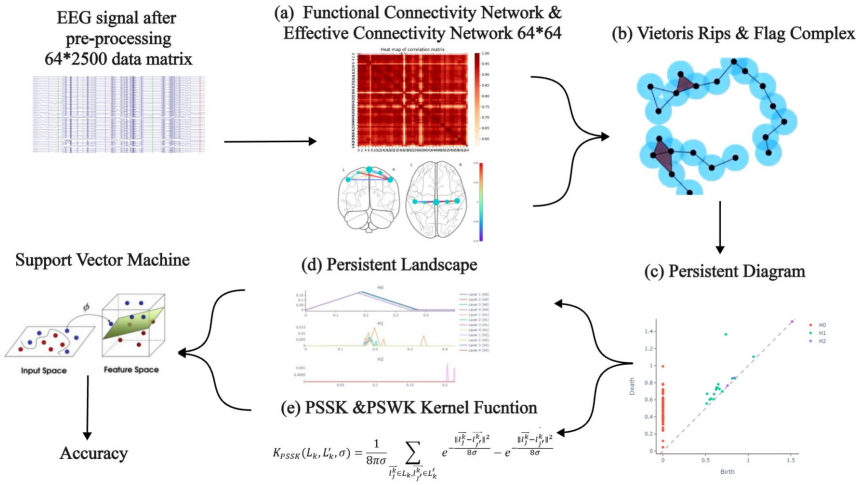


Fig.2: Process chart of data analysis. After preprocessing, data is used to extract (a) functional connectivity network (FC) to construct correlation matrix and effective connectivity networks (EC) to construct directed network; (b)(c) persistent diagrams are then constructed using Vietoris Rips and Flag complex to extract topological features from undirected/directed graph. Finally, topological features are input by building (d) persistent landscapes and (e) kernel functions into SVM machine learning for classification

2.4.1. Brain Connectivity Networks

Structural connectivity referred to the physical wiring between channels, whereas functional and effective connectivity networks derive from neuronal activity signals from a multidimensional time series [6]. In this paper, strength of channels' relationships used to distinguish between the kinds of movements made by subjects, where presented by functional and effective connectivity network.

Functional connectivity was generally decided by the coherence in different channels in EEG signals [7]. And correlation matrices were recognized as indicating linear connectivity, especially the Pearson correlation coefficient could be used to assess the linear interdependency between time series. Due to functional connectivity network only defined the presence of statistical dependencies between time series data. One of the major problems was it could not describe the causal relationships of functional interaction among the channels [7]. We also applied the effective connectivity network to remedy this issue.

Effective connectivity outlines the causal impacts that elements within a neural system impose on one another. This type of connectivity is measured using pairwise Granger causality [8]. Granger causality could be used as a parameter to assess whether a time series can predict another. By employing pairwise Granger causality, more robust estimates are obtained [9].

2.4.2. Topological Data Analysis

A simplicial complex is a topological object obtained by "gluing" together simple shapes such as points, links, surfaces, and higher dimension objects. For mathematical details, we refer to [10]. In short, simplicial complex is a generalization of a triangle to any dimension. For example, in 2-D, we refer to a 2-D simplex as a triangle, a 3-D simplex as a regular tetrahedron, and a 4-D simplex as a tetrahedron. In our project, we use Vietoris-Rips complexes and Flagser complexes, where each simplicial complexes corresponds to a subset of points whose supremum of distances between their pairs (diameter) is less than a given threshold, to construct abstract simplicial complexes from a set of points. The persistence diagram summarises information about the appearance and disappearance of topological features. VR complexes and Flagser complexes were used to form the persistence diagrams in the paper, where respectively correspond undirected graph and directed graph.

Here we will introduce the mathematic principle of the procedure and how to get persistent diagram (PD) from Vietoris-Rips complexes and Flagser complexes [11].

In a given dimension d, if there is an inclusion of one simplicial complex to another i: $K_1 \subseteq K_2$, then an inclusion map will exist on the d-chains $i : C_d(K_1) \rightarrow C_d(K_2)$ when recognizing small complex in any chain as one in the larger chain. And this can be extended to a map on homology $i : H_d(K_1) \rightarrow H_d(K_2)$ by sending [a] $\in H_d(K_1)$ to the class in $H_d(K_2)$ with the same representative. The persistent homology procedure can be defined as follow:

Firstly, we will give a filtration:

$$K_1 \subseteq K_2 \ldots \subseteq K_N \tag{5}$$

Then a sequence of maps on the homology will be obtained as:

$$H_d(K_1) \rightarrow H_d(K_2) \rightarrow \ldots \rightarrow H_d(K_N) \tag{6}$$

The definitions of birth and death are defined as: If $[\alpha]$ does not occur in the homology map $H_d(K_{i-1}) \rightarrow H_d(K_i)$, the class $[\alpha] \in H_d(K_i)$ will be defined to be born at i. The class dies at j if $[\alpha] \neq 0$ in $H_d(K_{j-1})$ but $[\alpha] = 0$ in $H_d(K_j)$.

PDs can be constructed to show the process of persistent homology after this definition is given. In a persistent diagram, each off-diagonal points $D = \{(b_1, d_1), \ldots, (b_k, d_k\}$, b_i and d_i represents that the homology is born at i and dies at j. The farther a point is from the diagonal, the longer time the class persists in the filtration and the stronger feature will be. The lifetime or persistence of a point $X = (b, d)$ in a persistence diagram D is given by pers $(x) = |b - d|$. Fig.3 plotted PHs of the sample movement 1 for subject 1 with the two brain network respectively. Each point represents a feature, whose x-axis is the filtering level at which the feature appears, and the y-axis is the filtering level where the feature disappears.

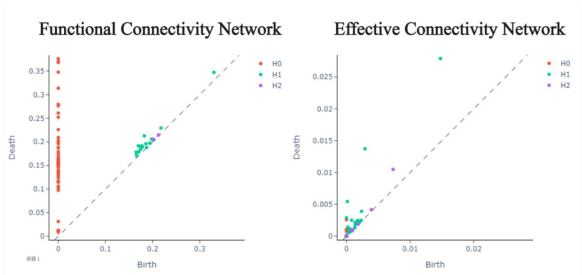


Fig.3: Persistent diagrams obtained from persistent diagrams of the movement 1 from subject 1

1. Persistent landscape

The persistent landscape is constructed by a series of persistent real-valued piecewise linear functions defined in the coordinates of the persistent diagram [12]. The piecewise function $f_{(b,d)}$ for each point (b, d) in the PD is defined. To quantify PLs, the distance of between a set of persistence landscapes is defined:

$$\| \lambda - \lambda' \|_\infty = sup_{k,t} | \lambda_k(t) - \lambda_k{}'(t) | \text{ where } \lambda_k(t) = k_{max}\{f_{(b,d)}(t)\} \quad (9)$$

where k_{max} represents the kth largest element. The persistent landscape is a sequence of functions where λ_k is called kth persistent landscape function. Fig.4 plotted PLs of the sample M1 for subject 1 with the two brain network respectively.

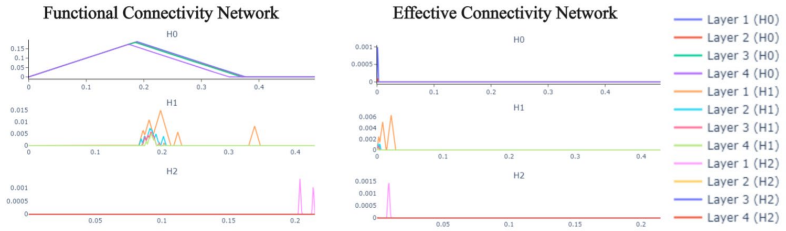


Fig.4: Persistent landscapes obtained from persistent diagrams of the movement 1 from subject 1

2. Persistent kernel function

From the topological perspective, various PH-based kernels have been also proposed [13]. In this paper, we apply the most frequently used kernel, Persistent Scale Space Kernel (PSSK), and Persistence Sliced Wasserstein Kernel (PSWK). The formulae for these two kernels are presented in the classification section below [13]. PSSK is defined as the inner product between the solutions of the heat diffusion equation, both originating from initial Dirac impulses positioned at the locations corresponding to the points of the persistent diagrams [14]; and for PSWK, which is one new kernel defined with Sliced Wasserstein metric. The metric slice the plane with lines, where the lines are to extrapolate the measures and integrate these distances over all possible lines [15].

Persistence diagrams have been proven useful in PSSK, as the first persistent homology related kernel function, but in practice, when using grid search, the bandwidth of the PSWK is easier to tune than the parameters of the PSSK [15]. So, we will take both kernels into account and we will find the one that is more suitable for our data.

$$K_{PSSK}(L_k, L'_k, \sigma) = \frac{1}{8\pi\sigma} \sum_{\vec{l_j^k} \in L_k, \vec{l_{j'}^k} \in L'_k} e^{-\frac{\|\vec{l_j^k} - \vec{l_{j'}^k}\|^2}{8\sigma}} - e^{-\frac{\|\vec{l_j^k} - l_{j'}^k\|^2}{8\sigma}} \quad (10)$$

Another PH-based kernel PSWK is defined by [15]

$$K_{SW}(Dg_1, Dg_2) \stackrel{\text{def}}{=} exp(-\frac{SW(Dg_1, Dg_2)}{2\sigma^2}) \quad (11)$$

where Dg_1, Dg_2 be two persistent diagrams

$$SW(Dg_1, Dg_2) \overset{\text{def}}{=} \frac{1}{2\pi}\int_{S1} w\left(\mu_1^\theta + \mu_{2\Delta}^\theta, \mu_2^\theta + \mu_{1\Delta}^\theta\right) d\theta \tag{12}$$

where $\mu_1^\theta = \sum_{p\in Dg_1} \delta_{\pi\theta(p)}$ and $\mu_{1\Delta}^\theta = \sum_{p\in Dg_1} \delta_{\pi\theta \circ \pi\Delta(p)}$

Following that, the results of PH were converted into PL and projected into higher dimensional space by kernel functions, that can be integrated into machine learning models support vector machine methods. The two parameters of models are σ and C values. C value is Penalty Coefficient of SVM model to get a better model for classification and finally choose the combination of two parameters with the highest accuracy. The key performance F1 score was used in evaluating classification models in the models, and for one run, the model randomly chose the 70% movements and train set, 30% movements as test set. And for each subject, we took 10 times run and calculated average as the final accuracy. Additionally, the accuracy presented by one-versus-one binary classification results with 3 types of motion conditions, 3 binary combinations.

3. Results

Multi-classification by the proposed model was performed to verify the feasibility of decoding coordinated directions of bimanual movements. For the accuracy result, $\alpha = 0.05$ was set in all significance test, and in all tables, classifier accuracy was presented as mean $\pm$ standard deviation (Std).

Table 1 presented the subject-independent accuracies of classifying with two different brain network and three different transformation methods of PH. The peak accuracy of decoding bimanual movement directions reached 72.48 ± 1.07 by functional connectivity network with PSSK function. Moreover, from Table 1, the two brain networks have no evident difference in accuracy, and the three subjects also have no significant difference, which outcome aligns with our anticipated outcomes.

And for the Table 2, blue color represents PL and orange color represents PSSK. Each accuracy presented as binary classification average percentage of the three subjects. the conclusion that movement 1 and movement 2 have higher accuracy, which means that these two types of movement were more differentiated than the other groups. And the results of Table 2 also showed no significant difference between the results of the two brain networks.

Table 1: Subject-independent Three Categories Classification with Brain Network.

Subject No.	Classification Accuracy (%)					
	Functional Connectivity Network			Effective Connectivity Network		
	Persistent Landscape	PSSK Function	PSWK Function	Persistent Landscape	PSSK Function	PSWK Function
Subject 1	68.42 ± 5.91	71.41 ± 4.19	67.99 ± 3.78	68.15 ± 6.17	70.84 ± 4.52	70.15 ± 4.10
Subject 2	65.61 ± 5.23	73.95 ± 4.86	65.75 ± 3.69	63.79 ± 4.52	72.64 ± 3.14	61.66 ± 5.82
Subject 3	65.34 ± 3.56	72.10 ± 3.12	63.65 ± 2.61	59.89 ± 2.19	70.11 ± 6.41	64.74 ± 4.22
Mean $\pm$ Std	66.05 ± 1.39	72.48 ± 1.07	65.80 ± 1.77	63.94 ± 3.37	71.20 ± 1.06	65.52 ± 3.51

Table 2: One-to-one Movement Binary Classification with Brain Network.

Functional	M1	M2	M3	Effective	M1	M2	M3
M1		69.55 ± 1.54	67.23 ± 2.51	M1		68.14 ± 2.61	65.75 ± 1.84
M2	73.19 ± 2.01		62.64 ± 1.68	M2	71.42 ± 3.40		63.14 ± 1.69
M3	66.85 ± 1.65	64.62 ± 0.94		M3	68.62 ± 2.08	62.74 ± 2.14	

Table 3: Result of real environment vs. virtual reality by functional connectivity network

Subject No.	REAL ENVIRONMENT			VIRTUAL REALITY		
	Persistent Landscape	PSSK Function	PSWK Function	Persistent Landscape	PSSK Function	PSWK Function
Subject 1	68.42±5.91	71.41±4.19	67.99±3.78	67.41±5.57	74.78±4.49	66.71±4.14
Subject 2	65.61±5.23	73.95±4.86	65.75±3.69	64.54±3.35	72.74±3.98	68.46±3.48
Subject 3	65.34±3.56	72.10±3.12	63.65±2.61	68.72±2.14	70.54±2.56	69.75±5.14
Mean±Std	66.05±1.39	72.48±1.07	65.80±1.77	66.89±1.75	72.69±1.73	68.31±1.25

4. Discussion

There are some possible reasons why the classification accuracy is not very high: 1) The current experimental paradigm design could make it challenging to distinguish the three bimanual movements. When performing the three types of bimanual movements, the brain activities may not differ significantly; 2) The number of participants in the experiment is relatively small, and individual differences among the subjects cannot be smoothed out. One subject only performs 150 movements, so these may not be sufficient for training in machine learning.

To achieve the further improvement, the directions were planned for in the future. First, we could refine the experimental paradigm and explore a range of bimanual movement combinations, whose movements should be used to demonstrate the differences in brain activity by simplifying common bimanual coordination movements in daily life; moreover, the existing methods could be improved and that new ones be tried. With enough computing power, we would try embedded point-cloud methods instead of neural networks. In addition, it would be beneficial to explore other complex structures, such as α-complexes and weighted Rips, in the construction of PH; finally, increasing the sample size of the experiment would be beneficial and the inclusion of subjects from different age groups and genders will enhance the generalizability of the experiments.

5. Conclusion

The paper provides a preliminary investigation into experimental paradigms for bimanual coordination. In this paper, we focus on classifying and decoding bimanual movements using PH-based machine learning. We first designed experiment for bimanual movements and collected EEG data from three subjects in real environments. We completed preprocessing of EEG data, then used methods combinations of TDA and machine learning for classification and obtained corresponding accuracy. Firstly, we have constructed functional connectivity network and effective connectivity network. Then, we used Vietoris-Rips and Flagser complexes to conduct persistent diagrams to capture the feature of EEG data. Finally, the persistent landscapes and persistent kernels were input in SVM for classification. The method with the best performance is to combine functional connectivity, PD and SVM with PSSK kernels, and the accuracy is 72.48 ± 1.07% in the real environment.

Acknowledgements

We thank professors Peng Chen, Rong Yin and Feng Gu, for their supports on equipment and data collection. We are also grateful to professor Chen and professor Gu's student teams for helping us complete the experiments.

References

[1] M. Zhang, Wu, J., Song, J., Fu, R., Ma, R., Jiang, Y. C., & Chen, Y. F. Decoding coordinated directions of bimanual movements from EEG signals. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 31, 248-259. https://doi.org/10.1109/TNSRE.2022.314056.

[2] L. R. Hochberg, Serruya, M. D., Friehs, G. M., Mukand, J. A., Saleh, M., Caplan, A. H., and Donoghue, J. P. (2006), Neuronal ensemble control of prosthetic devices by a human with tetraplegia, Nature, vol. 442, no. 7099, pp. 164–171, 2006.
[3] J. Wang, L. Bi, W. Fei, & C. Guan, Decoding Single-Hand and Both-Hand Movement Directions From Noninvasive Neural Signals, in *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 6, pp. 1932-1940, June 2021, doi: 10.1109/TBME.2020.3034112.
[4] L. K. Gallos, Makse, H. A., & Sigman, M. (2012). A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. Proceedings of the National Academy of Sciences, 109(8), 2825–2830. https://doi.org/10.1073/pnas.1106612109.
[5] Y. M. Zhang, J. Y. Yang, S. J. Xu and J. Y. Song, Z. Yang, F. S. Cheong, P. Chen, Comparative analysis of EEG signals in bimanual coordination: real vs. virtual environments for rehabilitation. Preprint.
[6] F. Takens. Detecting strange attractors in turbulence. In D. Rand & L. S. Young (Eds.), Dynamical Systems and Turbulence, Warwick 1980 (Lecture Notes in Mathematics, Vol. 898). Springer. https://doi.org/10.1007/BFb0091924.
[7] H.-J. Park, & K. Friston, (2013). Structural and functional brain networks: from connections to cognition. Science, 342(6158), 1238411. https://doi.org/10.1126/science.1238411.
[8] L. Barnett, & A. Seth, (2013). The MVGC multivariate Luigi Caputi causality toolbox: a new approach to Granger-causal inference. *Journal of Neuroscience Methods, 223*. https://doi.org/10.1016/j.jneumeth.2013.10.018.
[9] A. Seth, (2007). Granger causality. *Scholarpedia, 2*(7), Article 1667.
[10] A. Myers, A., Munch, E., & Khasawneh, F. A. (2019). Persistent homology of complex networks for dynamic state detection. Phys Rev E. 2019 Aug;100(2-1):022314. doi: 10.1103/PhysRevE.100.022314. PMID: 31574743.
[11] L. Caputi, Pidnebesna, A., & Hlinka, J. (2021). Promises and pitfalls of topological data analysis for brain connectivity analysis. NeuroImage vol. 238 (2021): 118245. doi: 10.1016/j.neuroimage.2021.118245.
[12] P. Bubenik, & P. Dłotko. A persistence landscapes toolbox for topological statistics. Journal of Symbolic Computation. Volume 78, January–February 2017, Pages 91-114. https://doi.org/10.1016/j.jsc.2016.03.009
[13] C. S. Pun, Xia, K., & Lee, S. X. (2018). Persistent-Homology-based machine learning and its applications--A survey. arXiv preprint arXiv:1811.00252.. Persistent-Homology-based machine learning and its applications—A survey. arXiv preprint arXiv:1811.00252.
[14] J. Reininghaus, S. Huber, U. Bauer, & R. Kwitt (2015). A stable multi-scale kernel for topological machine learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4741-4748.
[15] M. Carrière, M. Cuturi, & S. Oudot (2017). Sliced Wasserstein kernel for persistence diagrams. *In Proceedings of the 34th International Conference on Machine Learning,70* (17). JMLR.org, 664-673.