ORIGINAL ARTICLE

Optimal Multivariate Transfer Entropy to Determine Differences in Short and Long-Range EEG Connectivity in Children with ADHD and Healthy Children

Ali Ekhlasi [1*], Ali Motie Nasrabadi [2], Mohammadreza Mohammadi [3]

[1] *Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran*

[2] *Department of Biomedical Engineering, Faculty of Engineering, Shahed University, Tehran, Iran*

[3] *Psychiatry and Psychology Research Center, Roozbeh Hospital, Tehran University of Medical Sciences, Tehran, Iran*

*Corresponding Author: Ali Ekhlasi

Email: ali.ekhlasi@srbiau.ac.ir

Received: 22 April 2024 / Accepted: 25 August 2024

Abstract

Purpose: Investigating brain connectivity using Electroencephalogram (EEG) is a valuable method for studying mental disorders, such as Attention-Deficit/Hyperactivity Disorder (ADHD), and optimizing and developing measures of effective connectivity can provide new insights into differences in brain communication in such disorders. Multivariate Transfer Entropy (MuTE) is a measure of causal connectivity that quantifies the influence of multiple variables on each other in a system. In this study, the MuTE measure was modified by incorporating an interaction delay parameter in connectivity calculations to create a measure with self-prediction optimality, which we named MuTE_{SPO}.

Materials and Methods: We applied MuTE_{SPO} to investigate EEG effective connectivity in healthy and ADHD children performing an attention task across five frequency bands and to compare brain connectivity differences between the two groups using statistical analysis.

Results: Our analysis revealed that children with ADHD exhibited excessive short-distance connections in all frequency bands while healthy children demonstrated stronger long-range connections in the alpha and gamma frequency bands. Moreover, excessive short-distance connectivity was observed in the delta and theta frequency bands in all brain regions, as well as in the alpha, beta, and gamma frequency bands between the central and parietal regions in children with ADHD. These connectivity patterns may contribute to impaired attention functions by impeding effective information transmission and reducing information processing speed in the brains of children with ADHD.

Conclusion: Our analysis presents a novel methodology for measuring effective connectivity and elucidates the differences in EEG brain connectivity between children with ADHD and healthy children.

Keywords: Transfer Entropy; Effective Connectivity Analysis; Electroencephalogram; Attention-Deficit/Hyperactivity Disorder; Optimized Multivariate Transfer Entropy.

DOI: https://doi.org/10.18502/fbt.v12i2.18274

1. Introduction

The area of brain connectivity can be generally categorized into three groups: structural connectivity, functional connectivity, and effective connectivity. Structural connectivity focuses on the nerve fibers that establish connections between various brain regions. Functional connectivity deals with the statistical co-activations of various brain regions. Effective connectivity, on the other hand, pertains to the characterization of the causal interactions between different brain regions [1]. The examination of effective connectivity networks can assist neurologists in investigating how neural disorders such as Alzheimer's [2], Schizophrenia [3], Epilepsy [4, 5], Autism Spectrum Disorder (ASD) [6, 7], and Attention-Deficit/Hyperactivity Disorder (ADHD) [8, 9] affect the causality of the brain.

ADHD is a common neurodevelopmental disorder in children. To diagnose this disorder, experts examine the criteria of inattention and the criteria of hyperactivity and impulsivity in children [10]. These criteria and symptoms of ADHD are specified in the Diagnostic and Statistical Manual of Mental Disorders IV text edition (DSM-IV-TR) [11] and the International Classification of Diseases (ICD-10) [12]. Approximately 5-6% is the estimated occurrence rate of ADHD in children [13]. Global statistics indicate that ADHD is more common in boys than in girls [14]. The behavior pattern in a child with ADHD can appear in different situations (such as school and home) and lead to performance problems in social, educational, or work environments [11].

EEG (Electroencephalogram) analysis is a widely used method for examining brain activity in both healthy individuals and those with mental disorders. By measuring the electrical activity of the scalp generated by neuronal ensembles of the cerebral cortex with millisecond time resolution, EEG can be used to identify brain functions and classify subjects with disorders from healthy subjects [15]. Despite its limited ability to precisely identify the anatomical sources of neural activity, EEG remains extensively employed in research due to its accessibility, cost-effectiveness, and superior temporal resolution in comparison to other brain imaging techniques like functional Magnetic Resonance Imaging (fMRI) or Magnetoencephalography (MEG) [16]. With advanced signal processing techniques such as connectivity analysis, researchers can begin to identify the neural networks involved in specific cognitive processes or mental disorders.

Most studies examining EEG brain connectivity in individuals with ADHD have focused on determining functional connectivity, which measures the statistical association between different brain regions [17-22]. Sripida *et al.* have found evidence of distributed dysconnectivity within and between large-scale brain networks in individuals with ADHD [20]. Other studies have found an increase in interhemispheric and intrahemispheric coherence in the frontal and central regions of the brain in subjects with ADHD [19, 21, 22]. Clark and Barry conducted multiple studies that revealed individuals diagnosed with ADHD exhibit increased intrahemispheric coherence within the delta, theta, and beta frequency ranges at shorter to medium distances. Additionally, they observed reduced laterality in the theta band and reduced connectivity between the frontal hemisphere within the delta and theta frequency bands [17, 18]. However, many of these studies have focused on functional connectivity measures, and have not provided information about the causality of communication between brain regions.

Recent studies have used effective connectivity measures to demonstrate the causal relationships between different brain regions in mental disorders such as ADHD. Talebi *et al.* used the direct Directed Transfer Function (dDTF) measure and the newly introduced measure (nonlinear Causal Relationship Estimation by Artificial Neural Network- nCREANN) in the investigation of effective connectivity, which showed a decrease in the connection between the temporal/frontal and temporal/occipital regions, and an increase in the connection between the frontal/parietal regions in ADHD compared to healthy children [8]. In recent research in this field, the use of effective connectivity measures based on information theory has been considered. Unlike measures that calculate effective connectivity based on the mathematical model of data distribution (such as Granger Causality (GC) and Directed Transfer Function (DTF)), the information theory-based measures are model-free and use probability functions to estimate causal relationships. Transfer Entropy (TE) is a basic measure in this category and is widely

used to quantify the directional flow of information between time series.

TE was first introduced by Schreiber [23] as a way to quantify the amount of information that is transferred from one time series (the source) to another (the target). Specifically, it measures how much information about the future of the target time series can be gained by considering the past of the source time series, after accounting for the influence of the past of the target series itself. In neuroscience, TE has been used to investigate the flow of directed information between different brain regions, particularly in research related to ADHD patients. TE has also served as the foundation for the development of other measures, including Phase Transfer Entropy (PTE), which can provide additional insights into the directed interactions between different brain regions. Studies utilizing effective connectivity measures such as PTE have shown promise in investigating the flow of information between brain regions in both healthy individuals and those with ADHD. A study that employed PTE (phase transfer entropy) as a measure of effective connectivity discovered notable dissimilarities in the information transmission between distant brain regions in both healthy children and those with ADHD. These differences were especially evident within the theta frequency band [9, 24].

Other studies have utilized effective connectivity measures to construct brain networks and extract features for classification purposes. Abbas *et al.* constructed brain networks using TE and statistically compared graph features between two groups [25]. In two recent papers, we utilized EEG brain networks constructed by PTE to classify EEG signals between two groups. In the first paper, we reported 89.1% classification accuracy using PTE vectors and an Artificial Neural Network (ANN) [26]. In the second research, we reported 91.2% classification accuracy using graph features and a Bayesian classifier [27].

One limitation of conventional effective connectivity measures such as TE, PTE, DTF, and GC is that they are bivariate, meaning that they only consider the interaction between two time series. Bivariate measures may introduce indirect connectivity as direct connectivity, leading to weaker performance compared to multivariate measures. To address this limitation, multivariate effective connectivity measures have been developed that consider the interaction between multiple time series simultaneously. The multivariate version of TE (MuTE) was introduced by Montalto *et al.* in 2014 [28]. Although the MuTE has the advantage of being multivariable, it has limitations related to the setting of its parameters. By optimizing the parameters of this measure, its performance can be improved to more reliably calculate effective connectivity in multivariate systems.

In our research, we optimized the multivariate Transfer Entropy by incorporating the interaction delay between two time series. The resulting measure, called MuTE with Self-Prediction Optimization ($MuTE_{SPO}$) was used to study the effective connectivity between brain regions in children with ADHD and healthy children using EEG data. Specifically, we investigated statistical differences in effective connectivity between the two groups across five frequency bands.

The remainder of this paper is structured as follows: the materials and methods section will first introduce the EEG database and data preprocessing, followed by an explanation of the measures of TE, MuTE, and the optimized mode of MuTE ($MuTE_{SPO}$). Then, the construction of effective connectivity differential plots and statistical analysis will be described. In the third section, the results of the statistical analysis in five frequency bands will be presented. The fourth section will discuss and compare the obtained results with those of previous studies. Lastly, the conclusion will be presented in the fifth section.

2. Materials and Methods

2.1. The EEG Database

This study utilized a publicly accessible EEG database consisting of 121 children between the ages of 7 and 12, comprising both boys and girls. The database included 61 children diagnosed with ADHD and 60 children who were considered healthy [29]. The database used in this study followed the 10-20 international standard system for electrode placement. It included the following channels: Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, and Pz. To track eye movements, two electrodes

were positioned below and above the right eye. The EEG data were recorded at a sampling rate of 128 Hz.

The EEG recording protocol was designed based on attention disorder in children with ADHD. Based on this, every child should pay attention to the monitor screen on which images with cartoon characters are displayed. In order to maintain the child's focus on the screen, they were instructed to count the number of characters present. The accuracy of the child's response was not taken into account. The number of characters within each image varied randomly between 5 and 16. Additionally, the images were designed to be of sufficient size to allow for comfortable counting. Consequently, the completed database contained uninterrupted EEG data for each individual child, with the length of each recording varying depending on the child's performance during the task.

It should be noted that the database indicated that children with ADHD were diagnosed by an expert child psychologist based on DSM-IV [11]. The control group consisted of children who had no psychological disorders, history of epilepsy, drug abuse, or head injury, and their health was confirmed. The database publishers reported that all subjects participated in the experiment voluntarily, informed written consent was obtained from their parents, and the data recording process was approved by the Institutional Review Board (IRB) and Ethics Committee of Tehran University of Medical Sciences (TUMS). Further details about the EEG database used in this study can be found in reference [29].

2.2. EEG Pre-Processing

To extract meaningful results from the EEG data analysis, several pre-processing steps in the EEGLAB toolbox (version 14.1.1) were performed to remove artifacts and noise sources. First, a common average reference was applied to the EEG signals, and subsequently, a band-pass Finite Impulse Response (FIR) filter of 1-48 Hz was applied to the EEG signals. Next, the CleanLine plugin was used to remove the line noise. We then visually inspected the data over time to identify intervals where there was noise from electrode displacement in the EEG signal and removed these intervals while minimizing data loss. Then, the common reference average was removed from all channels in each subject. To remove eye blinks and muscle artifact components, we used Independent Component Analysis (ICA) and identified and removed the relevant components from the EEG signal. In the end, we utilized FIR filters with zero-phase shift to extract specific frequency bands - delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (30-45 Hz) - from the EEG signal.

These pre-processing steps allowed us to obtain high-quality EEG data that were suitable for further analysis. By removing the noise sources and artifacts, we were able to identify the EEG connectivity patterns and neural mechanisms underlying ADHD.

2.3. Bivariate Transfer Entropy (TE)

Bivariate Transfer Entropy (TE) is a model-free measure based on information theory that estimates the effective connectivity (linear and non-linear) between two processes of X and Y [23, 30]. It measures the amount of information that process X provides about the future behavior of Y, above and beyond the information that can be obtained from Y's own past behavior.

TE was obtained by reformulating Wiener's principle [31] in terms of conditional mutual information for Markov processes by Schindler [32] and Palus [33] as follows (Equation 1):

$$TE(X \to Y) = I(Y^+, X^- | Y^-) \quad (1)$$

The TE from X to Y in this Equation 1 is the mutual information between the past states of the processes X (X^-) and future states of the processes Y (Y^+), assuming that we know the past states of the Y (Y^-). This conditional mutual information function is rewritten with terms of Probability Mass Functions (PMFs) in the following form (Equation 2):

$$TE(X \to Y) = \sum p(Y^+, Y^-, X^-) Log \frac{p(Y^+ | Y^-, X^-)}{p(Y^+ | Y^-)} \quad (2)$$

Where the probability associated with the vector variable **a** is denoted as $p(\mathbf{a})$, while the probability of observing **b** given the values of **a**, represented by $p(\mathbf{b}|\mathbf{a})$, is calculated as $p(\mathbf{b}|\mathbf{a}) = p(\mathbf{a},\mathbf{b})/p(\mathbf{a})$. Considering Shannon entropy ($H(f) = -\sum p(f)\, Log\, p(f)$), the TE function can be

represented as the subtraction of two conditional Shannon entropies (Equation 3):

$$TE(X \to Y) = I(Y^+, X^- | Y^-) = H(X^- | Y^-)\, H(X^- | Y^+, Y^-) \quad (3)$$

Since Shannon entropy can be interpreted as the information in the system, Equation 3 can be explained as follows: The information transferred from X to Y is the average amount of information that we get from observing X^- after having already observed Y^- minus the information that is exclusively in X^- and not in Y^+ and Y^-. X^- and Y^- were constructed by determining the embedding delay (τ) and embedding dimension (d) of the state space of each system [30] (Equation 4):

$$X^- = x_t^d = (x_t, x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-(d-1)\tau})$$
$$Y^- = y_t^d = (y_t, y_{t-\tau}, y_{t-2\tau}, \ldots, y_{t-(d-1)\tau}) \quad (4)$$

If we write Y^+ as y_t, another representation of bivariate TE with the expansion of conditional Shannon entropies and embedding parameters is as follows (Equation 5):

$$TE\ (X \to Y) = H(x_t^d | y_t^d) - H(x_t^d | y_t, y_t^d) = H(x_t^d, y_t^d) - H(y_t^d) - H(x_t^d, y_t, y_t^d) + H(y_t, y_t^d) \quad (5)$$

In their study in 2013, Wibral *et al.* [34] highlighted that an underestimation of the information in the past values of (X^-) and past Y (Y^-) could result in a potentially high amount of information transfer from X to Y. To address this, they proposed an enhanced version of transfer entropy called TE_{SPO}. This improved method incorporated an interaction delay parameter (u) that satisfied the requirement of self-prediction optimality. Considering the interaction delay (u), Wibral *et al.* [34] rewrote the TE with interaction delay (u) (Equations 6, 7):

$$TE_{SPO}(X \to Y, u) = \sum_t p(y_t, x_{t-u}^d, y_t^d)\, Log\, \frac{p(y_t | y_t^d, x_{t-u}^d)}{p(y_t | y_t^d)} \quad (6)$$

$$TE_{SPO}(X \to Y, u) = H(x_{t-u}^d | y_t^d) - H(x_{t-u}^d | y_t, y_t^d) = H(x_{t-u}^d, y_t^d) - H(y_t^d) - H(x_{t-u}^d, y_t, y_t^d) + H(y_t, y_t^d) \quad (7)$$

In this formula, if u = 0, TE_{SPO} will be equal to the classic TE estimator. The performance of the TE_{SPO} estimator is shown in Figure 1 when the true interactive delay (δ) between two time series, X and Y, is known. Figure 1a. indicates X and Y time series with source and target titles, respectively, interacting from X to Y with the true interaction delay (δ). The

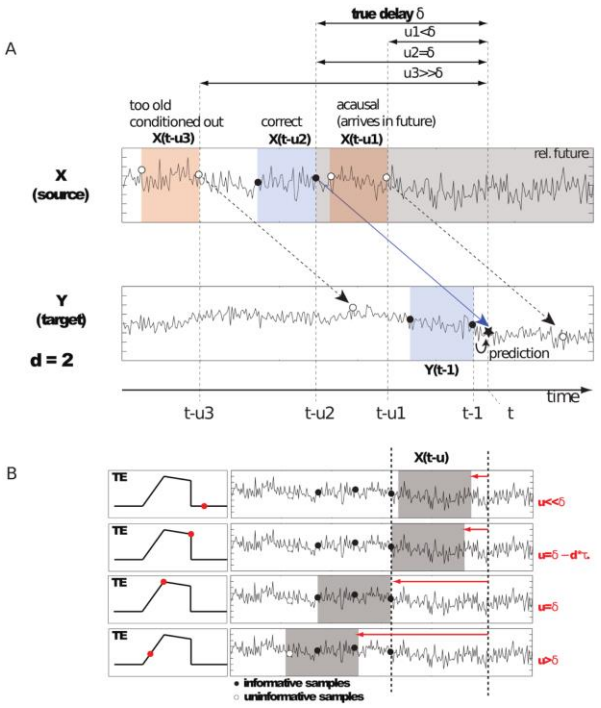

Figure 1. Central concepts underlying the reconstruction of interaction delay using the TE_{SPO} estimator [34]. Adapted from M. Wibral *et al.*, "Measuring information-transfer delays," *PloS one*, vol. 8, no. 2, p. e55809, 2013, Figure 1

star represents a sample of the Y (y_t) for which the transfer entropy function considers the X^- and Y^-. Circles (white and black) represent samples selected from X and Y as X^- and Y^-, respectively. In this Figure, embedding dimension (d) set to 2, and the embedding delay (τ) marked with colored boxes for both X^- and Y^-. The three values for interaction delay (u) are shown as follows: u1< δ, u2 = δ and u3>> δ. Figure 1b. shows how the TE changes for different u values. If the value of u corresponds to the true interaction delay value between two time series, so that $\mathbf{X}^-$ (gray boxes) contains informative samples, the TE estimator reaches its maximum value. Using TE_{SPO} the dominant information-transfer delay (δ) from X to Y can be estimated as in Equation 8 [34].

$$\delta = arg\ max_u\ TE_{SPO}(X \to Y, u) \quad (8)$$

When there is a single information transfer delay from X to Y, the parameter δ can be understood as the actual information transfer delay from X to Y. However, in cases where there are multiple coupling delays from X

to Y, Wibral *et al*. suggested averaging the TE_{SPO} over values over the range of delays u:[1 u_{max}][34]. This method has been used in other research studies as well [35].

2.4. Optimized Multivariate Transfer Entropy

In real-world systems, only two variables are rarely involved, and in most systems, there are multiple variables that interact with each other. The information interaction of other variables in a multivariate system affects the value of TE between X and Y. To address this limitation, multivariate TE (MuTE) has been introduced, which includes the information in the variables (X and Y) and other time series ($Z= Z^{(k)}_{k=1,...,M-2}$) in the TE calculations. Here, M is the number of variables in the system. The definition of multivariate transfer entropy from X to Y, conditioned on Z, is as follows (Equation 9):

$$MuTE(X \rightarrow Y|Z) = \sum_{t} p(Y^{+}, Y^{-}, X^{-}, Z^{-}) \, Log \, \frac{p(Y^{+}| \, Y^{-}, X^{-}, Z^{-})}{p(Y^{+}| \, Y^{-}, Z^{-})} \quad (9)$$

MuTE calculates the transfer of information from X to Y while already knowing the information in Y^{-} and Z^{-} [28]:

$$MuTE(X \rightarrow Y|Z) = I(X^{-}, Y^{+} | \, Y^{-}, Z^{-}) \quad (10)$$

The conditional mutual information function can be represented using conditional entropy, as demonstrated by Equations 5 and 7 [36, 37]:

$$\begin{aligned} MuTE(X \rightarrow Y|Z) &= H(x_t^d | \, y_t^d, z_t^d) \\ &\quad - H(x_t^d | \, y_t, y_t^d, z_t^d) \\ &= H(x_t^d, y_t^d, z_t^d) - H(y_t, y_t^d) \\ &\quad - H(x_t^d, y_t, y_t^d, z_t^d) \\ &\quad + H(y_t, y_t^d, z_t^d) \end{aligned} \quad (11)$$

In this study, we modified the MuTE by entering the interaction delay (u) into the MuTE equation. Similar to Wibral's definition of TE_{SPO}, our modified measure was named $MuTE_{SPO}$ (Multivariate transfer Entropy with self-prediction optimality). In the modified mode, $MuTE_{SPO}$ was calculated in the time interval from 1 to u_{max}, so that the time required for the interaction delay is considered in the $MuTE_{SPO}$ calculation. The equation of the $MuTE_{SPO}$ can be shown in the following order (Equation 12):

$$\begin{aligned} MuTE(X \rightarrow Y|Z,u) &= H(x_{t-u}^d | \, y_t^d, z_t^d) \\ &\quad - H(x_{t-u}^d | \, y_t, y_t^d, z_t^d) \\ &= H(x_{t-u}^d, y_t^d, z_t^d) - H(y_t^d, z_t^d) \\ &\quad - H(x_{t-u}^d, y_t, y_t^d, z_t^d) \\ &\quad + H(y_t, y_t^d, z_t^d) \end{aligned} \quad (12)$$

And in this way, the dominant information-transfer delay (δ') by $MuTE_{SPO}$ can be defined as (Equation 13):

$$\delta' = arg \, max_u \, MuTE_{SPO}(X \rightarrow Y|Z, u) \quad (13)$$

Assuming the value of u to be zero in Equation 12 results in the $MuTE_{SPO}$ function being equivalent to the classic MuTE function.

2.5. Construction of Directed Differential Connectivity Graph (dDCG)

In this study, we first estimated the directional connectivity strength between each pair of channels for each subject using the $MuTE_{SPO}$ measure. To achieve our goal, we segmented each subject's EEG signals into non-overlapping windows of 8 seconds (1024 samples). Next, we computed the $MuTE_{SPO}$ value for each pair of channels within each segment. By averaging the values obtained for each subject in each frequency band across the segments, we derived a distinct $MuTE_{SPO}$ value. Our previous research demonstrated that a minimum length of 1024 samples is required to obtain reliable results from TE-based measures in time series [38]. Given the assumption of stationarity in EEG signals, we also used a window length of 1024 samples in this study. Then, a statistical test between two groups was performed for each directed link. During the statistical test, we examined the null hypothesis, which states that the strength of each link remained the same across both groups. Conversely, the alternative hypothesis posits that there were significant differences in the strength of the link between the two groups. For the statistical analysis, we utilized a t-test (Equation 14) to compare the average strength of each link between the two groups. In this equation, μ_1 and μ_2 represent the sample mean of link strength for ADHD and healthy groups, while σ_1 and σ_2 represent the sample variances of link strength for ADHD and healthy groups, respectively. The number of subjects in the healthy

group is denoted as N1, and the number of subjects in the ADHD group is denoted as N2.

$$t = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}}} \quad (14)$$

To test the null hypothesis, the t null distribution was estimated with non-parametric permutation tests for each link (N_p: Number of Permutation= 5000). Subsequently, the significance of the t-statistic for each link was evaluated by comparing its actual value with the null distribution obtained. This comparison was done using the p-value. To identify significant differences between the two groups, FDR-corrected p-values were employed. These corrected p-values were considered significant at a threshold of <0.01. These significant differences represent directed differential connectivity graphs (dDCG) for each of the five frequency bands.

3. Results

In this study, we introduced the $MuTE_{SPO}$, which calculates the optimized transfer entropy in the case of a multivariate system. We used this measure to determine the brain-effective connectivity of healthy and ADHD children based on their 19-channel EEG signals. In order to examine the connectivity patterns in both individuals with ADHD and healthy individuals, we computed the $MuTE_{SPO}$ between each pair of channels across five frequency bands. This analysis encompassed all subjects and channels in the study. To calculate the $MuTE_{SPO}$ measure, we modified the TIM 1.3.0 toolbox to convert the standard "MuTE" measure to optimized MuTE measure by taking into account the interaction delay (u). The TIM toolbox is a cross-platform C++/Matlab library for efficient non-linear time-series analysis. The TIM toolbox calculates the Shannon entropies (Equations 11 and 12) using the k-Nearest Neighbor (kNN) estimator [39]. Alternative Shannon entropy estimators, such as binning and kernel methods, can introduce biases that impact the precision of the estimated direction of information flow [40]. Moreover, these estimators may occasionally produce reversed estimates of information flow direction, despite being faster compared to the K-Nearest Neighbor (K-NN) estimator [41].

Due to the fact that the calculation of TE and its variants, including $MuTE_{SPO}$, "MuTE," $MuTE_{SPO}$, involves four terms of Shannon entropy, each with different dimensions, choosing a fixed value for K can lead to bias error in the calculation of entropy and the distance between points [40]. To tackle this concern, Kraskov *et al.* suggested choosing the value of k in the highest dimensional space as a starting point, and subsequently projecting the computed distance to lower dimensional spaces to determine the range for searching neighbors [40]. In our study, we opted to set the value of K as 10. This decision was based on prior research that demonstrated the K-Nearest Neighbors (K-NN) estimator's performance remains fairly stable even when this parameter is subject to reasonable variations [40, 42].

There are two different methods to determine the optimal delay (τ) and the optimal dimensions (d) of the embedding. The first method, the Ragwitz method, is more suitable for high-dimensional multivariate systems and works by minimizing the prediction error [43]. This method calculates (τ) and d simultaneously and was chosen in our research because the $MuTE_{SPO}$ measure is inherently multivariate, and we used it to determine 19-channel EEG connectivity. The second approach, which is suitable for low-dimensional systems, involves utilizing the first local minimum of the auto-mutual information to determine the value of τ. Then, the Cao criterion [44] is employed to estimate the embedding dimension (d). In order to implement the Ragwitz method, we explored a range of 1 to 10 samples for the embedding delay (τ) and 1 to 10 for the embedding dimension (d). For each EEG channel of every subject, we computed the mean squared prediction error ($e^2(\tau, d)$) individually. The optimal values of τ and d for the entire dataset were selected by minimizing $e^2(\tau, d)$ across all channels and subjects. We set the interaction delay $u_{max}$to 10 samples, which indicated the use of a 10-sample delay for transferring information between two signals from one channel to another channel using $MuTE_{SPO}$. The unique value for $MuTE_{SPO}$ was obtained by averaging the values for $MuTE_{SPO}$ across u values ranging from 1 to 10.

We then used statistical analysis to construct dDCG for each frequency band between the two groups, as

shown in Figure 2(a-e). The Figures depict red lines to represent connections that exhibit statistically significant differences between the two groups, with higher average strength observed in individuals with ADHD. On the other hand, blue lines indicate connections with statistically significant differences between the two groups, where the average strength is higher in healthy individuals. Overall, these Figures provide a visual representation of the connectivity patterns observed in individuals with ADHD compared to healthy individuals in different frequency bands.

The results of the study indicate that individuals with ADHD exhibit unique connectivity patterns in different frequency bands. Upon a general examination of Figure 2, it becomes apparent that the delta and theta frequency bands exhibit a significantly higher number of red arrows, indicating stronger connectivity, in children with ADHD compared to healthy children. Interestingly, as the frequency range increases, the number of red arrows decreases. In the alpha, beta, and gamma frequency bands, nearly all the red arrows were observed over short distances between electrodes. Additionally, in certain frequency bands such as theta, alpha, and gamma, blue arrows were observed, indicating stronger long-range effective connections in healthy children as opposed to those with ADHD.

4. Discussion

In this study, we introduce the optimized multivariate transfer entropy, termed $MuTE_{SPO}$, which represents a significant advancement in quantifying effective connectivity. Unlike traditional multivariate transfer entropy methods, which typically calculate connectivity based on a fixed interaction delay, $MuTE_{SPO}$ incorporates a range of interaction delays into its analysis. This innovation addresses a critical limitation of classic multivariate transfer entropy, where using a single, fixed delay might miss important causal interactions if the true delay between variables falls outside the chosen delay window. By averaging the MuTE values over a range of interaction delays, $MuTE_{SPO}$ provides a more accurate and comprehensive measure of connectivity, reducing the risk of overlooking significant causal relationships that could be missed with traditional methods.

$MuTE_{SPO}$ was employed to evaluate effective connectivity in the brains of children, both those who were healthy and those diagnosed with ADHD, using their EEG signals while engaged in an attention task. A statistical analysis was conducted to assess the difference in effective connectivity between the two groups. The results of our study revealed that children diagnosed with ADHD exhibited excessive short-range connections, particularly at low frequencies, and impaired long-range connections at high frequencies, in comparison to their healthy subjects.

The statistical test conducted on the $MuTE_{SPO}$ values in Figure 2 indicates stronger brain-effective connections at short intervals in children with ADHD across all frequency bands (red arrows). Notably, In the delta frequency band, individuals with ADHD have a significantly higher number of stronger connections in both short and long connections compared to healthy individuals (Figure 2a). However, as the frequency increases, the number of short stronger directed brain connections of children with ADHD decreases. Our finding regarding the potential role of delta frequency band connectivity in the difference between the two groups is in line with prior research. Luo *et al.* conducted a frequency analysis and found that the fronto-central region exhibited higher delta power in their study [43]. In our previous study, we employed directed PTE within the delta frequency range and discovered a noteworthy distinction between the two groups concerning internal connections in the anterior region. The ADHD group exhibited higher values compared to the other group, indicating a significant difference [9]. Two recent studies have reported increased nodal efficiency in the delta frequency band in individuals with ADHD [27, 46].

Based on statistically significant distinctions between the two groups, our findings in the theta frequency range indicated that the ADHD group exhibited a reduced number of stronger effective connections over shorter distances compared to the delta band. However, these connections were still higher in number compared to the alpha frequency band. Additionally, we observed a few long-range connections between brain regions in a theta frequency band, with significantly higher strength in healthy individuals compared to those with ADHD (blue arrows). Figure 2(a and b) demonstrates that the

delta and theta frequency bands exhibit a greater number of significant directed connections between the two groups compared to the other frequency bands.

In these two frequency bands, the stronger short connections in children with ADHD were distributed across all brain regions. The crucial role of connectivity in the delta and theta frequency bands in distinguishing and classifying between the ADHD and healthy groups has been supported by numerous studies. Ahmadlou discovered that utilizing criteria like Synchronization Likelihood (SL) and Fuzzy Synchronization Likelihood (FSL) to assess functional connectivity allowed for the identification of synchronization patterns achieving classification accuracies of 87.5% and 95.6% in the theta and delta frequency bands, respectively [47, 48]. Our last two studies on classifying between two groups using features extracted from the PTE connectivity graph showed that the highest classification accuracy was achieved using features from the delta and theta frequency bands.

Based on the data presented in Figure 2c, it was observed that children with ADHD showed stronger effective connections primarily over short distances in the alpha band, which is consistent with previous findings in other frequency bands. These connections were mostly between central regions and between central and parietal regions. Furthermore, in this frequency band, the statistical analysis of $MuTE_{SPO}$ revealed the presence of long-range effective connections that were stronger in healthy children compared to those with ADHD. Specifically, these directed connections were frequently observed between the occipital region of the right hemisphere and the frontal and prefrontal regions.

After analyzing the $MuTE_{SPO}$ values in the beta frequency band, our results indicate that the mean power of all the connections with short distances that exhibited a statistically significant difference between the two groups was higher in children with ADHD (Figure 2d). We found that the stronger connections were primarily located between the central and parietal regions (except F3→F7). Through our analysis, we discovered that four pairs of electrode connections located in these regions displayed statistically significant differences between the two groups in both directions (C3↔P3, Cz↔P4, C4↔P4, T4↔T6). Two pairs of these effective connections located between the central and parietal regions of the right hemisphere (Cz↔P4 and C4↔P4) in the gamma frequency band also had significant strength in the ADHD group. Also, in the last frequency band, as shown by the blue arrows in Figure 2e, stronger connections were identified in healthy children, especially at long distances between electrodes.

Based on our analysis of the statistical results obtained from the $MuTE_{SPO}$ for effective connectivity between two groups of children (ADHD and healthy), we have drawn the following conclusions: Children diagnosed with ADHD demonstrated heightened effective connectivity between brain regions in the low-frequency bands (delta and theta) when compared to healthy children. These increased connections were frequently observed over shorter distances and between electrodes that were in close proximity to each other. However, as the frequency increased, the number of stronger connections decreased in children with ADHD. Conversely, long-range connections were observed in high-frequency bands (alpha and gamma) that were statistically stronger in healthy children than in children with ADHD.

These findings suggest that the presence of short and extensive connections in children with ADHD may limit the transmission of information within their brain regions, causing information to become trapped over sub-networks and reducing the speed of information transfer. This limitation in information transmission was observed in different frequency bands and specifically between central and parietal regions in alpha, beta, and gamma. In contrast, the stronger long-range connections of healthy children allow for rapid information transfer. Therefore, the excessive short connections and impaired long-range connections in the brain networks of children with ADHD may contribute to attention and behavioral disorders in these individuals. Our research results validate previous studies that utilized graph theory to examine EEG signals and fMRI data and also indicate new insights to the existing knowledge. Earlier studies revealed that the brain networks of individuals with ADHD have greater levels of separation and less integration [27, 49, 50], and the research conducted on the interaction between MEG channels also showed that the connectivity for ADHD patients was more than that of control subjects in all frequency bands and in short distances [51]. In addition to validating prior

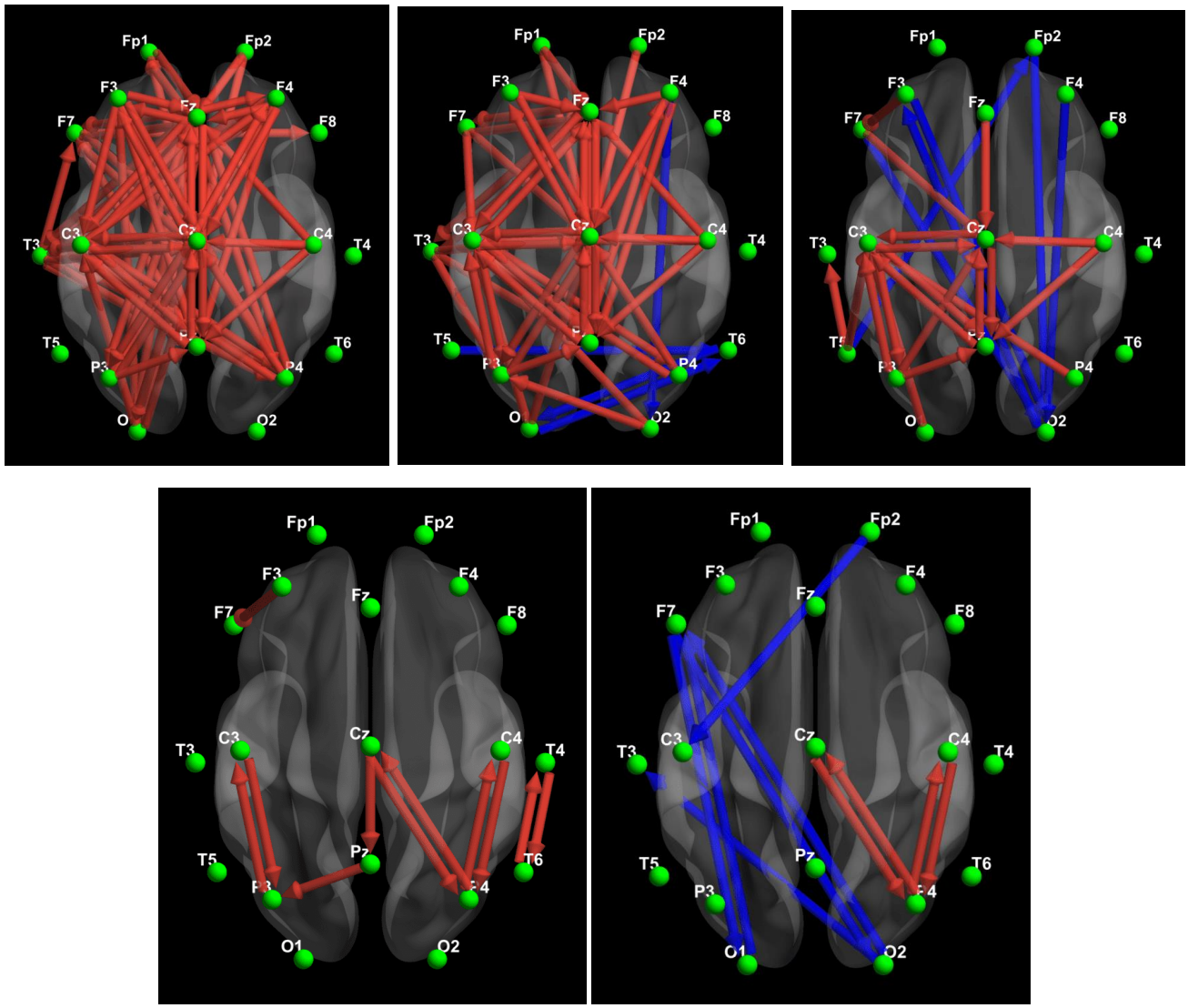

Figure 2. dDCG between ADHD subjects and healthy subject in five frequency bands. (a) Delta band, (b) Theta band, (c) Alpha band, (d) Beta band, (e) Gamma band. The red arrows represent significant differences connections with higher average strength in subjects with ADHD, while the blue arrows represent significant differences connections with higher average strength in healthy subjects

research findings, our study demonstrated that information trapping within brain sub-networks in individuals with ADHD is more pronounced at lower frequencies and diminishes at higher frequencies. Specifically, this trapping predominantly occurs in the connectivity between the central and parietal regions of the brain.

Besides introducing the $MuTE_{SPO}$ measure and the observed differences in EEG connectivity between healthy children and children with ADHD using this modified connectivity criterion, this study has some limitations. One limitation is the number of EEG channels in the database, which consists of 19 channels. For a more comprehensive understanding of the connectivity differences between these two groups, incorporating a greater number of EEG channels would be beneficial. Another limitation is that our investigation of brain connectivity was conducted through scalp electrodes. To better understand the source of connectivity in the brain, using source localization methods with a higher density of EEG channels is recommended, as it can help address volume conduction issues more effectively. In our study, we attempted to mitigate volume conduction effects and ensure accurate connectivity findings by employing a combination of preprocessing techniques, including average reference and ICA for spatial filtering, along with surrogate data analysis to validate the robustness of our results.

It should also be noted that, similar to Transfer Entropy (TE), the $MuTE_{SPO}$ measure, which is model-free, requires more computational processing to

identify causal connectivity compared to other measures that use data distribution models, such as Granger Causality (GC) and Directed Transfer Function (dDTF). Additionally, because $MuTE_{SPO}$ investigates all pairwise EEG connections due to its multivariate nature, it requires more computation time compared to bivariate measures like TE, which can misinterpret indirect connections as direct. Despite its higher computational cost, $MuTE_{SPO}$ has three advantages: it is model-free, thus eliminating modeling errors; it is multivariate, avoiding the misinterpretation of indirect connections as direct; and it considers a range of interaction delays with self-prediction optimality, minimizing the risk of missing interactions.

5. Conclusion

In conclusion, we present an enhanced version of multivariate transfer entropy, termed $MuTE_{SPO}$ which enabled the assessment of both direct and indirect information flow between variables while considering the influence of all other variables within the multivariate system. The $MuTE_{SPO}$ incorporates an interaction delay parameter to measure prominent information transfer between two variables. Our statistical analysis of $MuTE_{SPO}$ revealed that children with ADHD exhibited excessive short-distance connectivity across all frequency bands, in contrast to healthy controls, who showed more robust long-range connections, particularly in the alpha and gamma bands. Furthermore, the observed excessive short-distance connectivity in ADHD was prevalent in delta and theta bands across all brain regions, and in alpha, beta, and gamma bands specifically between central and parietal regions. This suggests a more pronounced information trapping within brain sub-networks in ADHD individuals, predominantly affecting lower frequencies and diminishing at higher frequencies, with notable effects in central-parietal connectivity. Future research should consider incorporating higher EEG channel numbers to determine the performance of $MuTE_{SPO}$ and to gain a more comprehensive understanding of EEG connectivity in ADHD. Additionally, employing source localization methods could provide insights into the origins of these connectivity patterns. To facilitate broader application and further investigation, developing a toolbox for $MuTE_{SPO}$" would be beneficial for the signal processing research.

References

1- Karl J Friston, "Functional and effective connectivity: a review." *Brain connectivity,* Vol. 1 (No. 1), pp. 13-36, (2011).

2- Jiali Huang, Jae-Yoon Jung, and Chang S Nam, "Estimating effective connectivity in Alzheimer's disease progression: A dynamic causal modeling study." *Frontiers in Human Neuroscience,* (2022).

3- Andreea Oliviana Diaconescu *et al.*, "Aberrant effective connectivity in schizophrenia patients during appetitive conditioning." *Frontiers in Human Neuroscience,* Vol. 4p. 239, (2011).

4- Nader Moharamzadeh and Ali Motie Nasrabadi, "A fuzzy sensitivity analysis approach to estimate brain effective connectivity and its application to epileptic seizure detection." *Biomedical Engineering/Biomedizinische Technik,* Vol. 67 (No. 1), pp. 19-32, (2022).

5- Shahdoust N, Cowan RL, Price TA, Kundu B, Davis TS, Rolston JD, *et al.* "Interictal Epileptiform Discharges Disrupt Neural Computations Underlying Cognitive Control and Value-based Decision Making.", (2024).

6- Xinling Geng *et al.*, "Abnormalities of EEG Functional Connectivity and Effective Connectivity in Children with Autism Spectrum Disorder." *Brain Sciences,* Vol. 13 (No. 1), p. 130, (2023).

7- Fatemeh Salehi, Mehrad Jaloli, Robert Coben, and Ali Motie Nasrabadi, "Estimating brain effective connectivity from EEG signals of patients with autism disorder and healthy individuals by reducing volume conduction effect." *Cognitive Neurodynamics,* Vol. 16 (No. 3), pp. 519-29, (2022).

8- Nasibeh Talebi and Ali Motie Nasrabadi, "Investigating the discrimination of linear and nonlinear effective connectivity patterns of EEG signals in children with Attention-Deficit/Hyperactivity Disorder and Typically Developing children." *Computers in Biology and Medicine,* Vol. 148p. 105791, (2022).

9- Ali Ekhlasi, Ali Motie Nasrabadi, and Mohammad Reza Mohammadi, "Direction of information flow between brain regions in ADHD and healthy children based on EEG by using directed phase transfer entropy." *Cognitive Neurodynamics,* Vol. 15 (No. 6), pp. 975-86, (2021).

10- Paul H Wender, "Attention-deficit hyperactivity disorder in adults." *Psychiatric Clinics of North America,* Vol. 21 (No. 4), pp. 761-74, (1998).

11- Ap Association, "Diagnostic and statistical manual of mental disorders (DSM-IV-TR), text revision." *(No Title),* (2000).

12- World Health Organization, The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research. *World Health Organization*, (1993).

13- Philip Asherson, Jan Buitelaar, Stephen V Faraone, and Luis A Rohde, "Adult attention-deficit hyperactivity disorder: key conceptual issues." *The Lancet Psychiatry*, Vol. 3 (No. 6), pp. 568-78, (2016).

14- RC Kessler *et al.*, "The effects of temporally secondary co-morbid mental disorders on the associations of DSM-IV ADHD with adverse outcomes in the US National Comorbidity Survey Replication Adolescent Supplement (NCS-A)." *Psychological medicine*, Vol. 44 (No. 8), pp. 1779-92, (2014).

15- Paul L Nunez and Ramesh Srinivasan, Electric fields of the brain: the neurophysics of EEG. *Oxford University Press, USA*, (2006).

16- Julián J González, Guzmán Alba, Soledad Mañas, Almudena González, and Ernesto Pereda, "Assessment of ADHD through electroencephalographic measures of functional connectivity." *ADHD-New Dir. Diagnosis Treat*, pp. 35-54, (2017).

17- Robert J Barry and Adam R Clarke, "Resting state brain oscillations and symptom profiles in attention deficit/hyperactivity disorder." in *Supplements to clinical neurophysiology*, Vol. 62: *Elsevier*, (2013), pp. 275-87.

18- Robert J Barry, Adam R Clarke, and Stuart J Johnstone, "A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography." *Clinical neurophysiology*, Vol. 114 (No. 2), pp. 171-83, (2003).

19- Franca E Dupuy, Adam R Clarke, Robert J Barry, Rory McCarthy, and Mark Selikowitz, "EEG coherence in children with attention-deficit/hyperactivity disorder: differences between good and poor responders to methylphenidate." *Psychiatry Research*, Vol. 180 (No. 2-3), pp. 114-19, (2010).

20- Chandra Sripada, Daniel Kessler, Yu Fang, Robert C Welsh, Krishan Prem Kumar, and Michael Angstadt, "Disrupted network architecture of the resting brain in attention-deficit/hyperactivity disorder." *Human brain mapping*, Vol. 35 (No. 9), pp. 4693-705, (2014).

21- Robert J Chabot, Henry Merkin, Lisa M Wood, Tracey L Davenport, and Gordon Serfontein, "Sensitivity and specificity of QEEG in children with attention deficit or specific developmental learning disorders." *Clinical Electroencephalography*, Vol. 27 (No. 1), pp. 26-34, (1996).

22- Robert J Chabot and Gordon Serfontein, "Quantitative electroencephalographic profiles of children with attention deficit disorder." *Biological psychiatry*, Vol. 40 (No. 10), pp. 951-63, (1996).

23- Thomas Schreiber, "Measuring information transfer." *Physical review letters*, Vol. 85 (No. 2), p. 461, (2000).

24- Ekhlasi A, Nasrabadi AM, Mohammadi MR. Analysis of effective connectivity strength in children with attention deficit hyperactivity disorder using phase transfer entropy. Iranian Journal of Psychiatry. 2021;16(4):374.

25- Ali Kareem Abbas, Ghasem Azemi, Sajad Amiri, Samin Ravanshadi, and Amir Omidvarnia, "Effective connectivity in brain networks estimated using EEG signals is altered in children with ADHD." *Computers in Biology and Medicine*, Vol. 134p. 104515, (2021).

26- Ali Ekhlasi, Ali Motie Nasrabadi, and Mohammadreza Mohammadi, "Classification of the children with ADHD and healthy children based on the directed phase transfer entropy of EEG signals." *Frontiers in Biomedical Technologies*, Vol. 8 (No. 2), pp. 115-22, (2021).

27- Ali Ekhlasi, Ali Motie Nasrabadi, and Mohammadreza Mohammadi, "Analysis of EEG brain connectivity of children with ADHD using graph theory and directional information transfer." *Biomedical Engineering/Biomedizinische Technik*, Vol. 68 (No. 2), pp. 133-46, (2023).

28- Alessandro Montalto, Luca Faes, and Daniele Marinazzo, "MuTE: a MATLAB toolbox to compare established and novel estimators of the multivariate transfer entropy." *PloS one*, Vol. 9 (No. 10), p. e109462, (2014).

29- Ali Motie Nasrabadi, Armin Allahverdy, Mehdi Samavati, and Mohammad Reza Mohammadi, "EEG data for ADHD/Control children." *IEEE Dataport*, (2020).

30- Michael Wibral, Raul Vicente, and Joseph T Lizier, Directed information measures in neuroscience. *Springer*, (2014).

31- Norbert Wiener, "The theory of prediction." *Modern mathematics for engineers*, (1956).

32- Katerina Hlaváčková-Schindler, Milan Paluš, Martin Vejmelka, and Joydeep Bhattacharya, "Causality detection based on information-theoretic approaches in time series analysis." *Physics Reports*, Vol. 441 (No. 1), pp. 1-46, (2007).

33- Milan Paluš, Vladimír Komárek, Zbyněk Hrnčíř, and Katalin Štěrbová, "Synchronization as adjustment of information rates: Detection from bivariate time series." *Physical Review E*, Vol. 63 (No. 4), p. 046211, (2001).

34- Michael Wibral *et al.*, "Measuring information-transfer delays." *PloS one*, Vol. 8 (No. 2), p. e55809, (2013).

35- Ali Khadem, Gholam-Ali Hossein-Zadeh, and Anahita Khorrami, "Long-range reduced predictive information transfers of autistic youths in EEG sensor-space during face processing." *Brain topography*, Vol. 29pp. 283-95, (2016).

36- Angeliki Papana, Catherine Kyrtsou, Dimitris Kugiumtzis, and Cees Diks, "Detecting causality in non-stationary time series using partial symbolic transfer

entropy: Evidence in financial data." *Computational economics*, Vol. 47 (No. 3), pp. 341-65, (2016).

37- Dimitris Kugiumtzis, "Partial transfer entropy on rank vectors." *The European Physical Journal Special Topics*, Vol. 222 (No. 2), pp. 401-20, (2013).

38- Ali Ekhlasi, Ali Motie Nasrabadi, and Mohammad Reza Mohammadi, "Improving Transfer Entropy and Partial Transfer Entropy for Relative Detection of Effective Connectivity Strength between Time Series." *Available at SSRN 4388493*.

39- Germán Gómez-Herrero, Wei Wu, Kalle Rutanen, Miguel C Soriano, Gordon Pipa, and Raul Vicente, "Assessing coupling dynamics from an ensemble of time series." *Entropy*, Vol. 17 (No. 4), pp. 1958-70, (2015).

40- Alexander Kraskov, Harald Stögbauer, and Peter Grassberger, "Estimating mutual information." *Physical Review E*, Vol. 69 (No. 6), p. 066138, (2004).

41- Daniel W Hahs and Shawn D Pethel, "Distinguishing anticipation from causality: Anticipatory bias in the estimation of information flow." *Physical review letters*, Vol. 107 (No. 12), p. 128701, (2011).

42- Daniele Marinazzo, Mario Pellicoro, and Sebastiano Stramaglia, "Causal information approach to partial conditioning in multivariate data sets." *Computational and mathematical methods in medicine*, Vol. 2012(2012).

43- Mario Ragwitz and Holger Kantz, "Markov models from data by simple nonlinear time series predictors in delay embedding spaces." *Physical Review E*, Vol. 65 (No. 5), p. 056201, (2002).

44- Liangyue Cao, "Practical method for determining the minimum embedding dimension of a scalar time series." *Physica D: Nonlinear Phenomena*, Vol. 110 (No. 1-2), pp. 43-50, (1997).

45- Na Luo *et al.*, "Aberrant brain dynamics and spectral power in children with ADHD and its subtypes." *European Child & Adolescent Psychiatry*, pp. 1-12, (2022).

46- Mingkan Shen, Peng Wen, Bo Song, and Yan Li, "ADHD children identification based on EEG using effective connectivity techniques." in *Health Information Science: 10th International Conference, HIS 2021, Melbourne, VIC, Australia, October 25–28, 2021, Proceedings 10*, (2021): *Springer*, pp. 71-81.

47- Mehran Ahmadlou and Hojjat Adeli, "Wavelet-synchronization methodology: a new approach for EEG-based diagnosis of ADHD." *Clinical EEG and neuroscience*, Vol. 41 (No. 1), pp. 1-10, (2010).

48- Mehran Ahmadlou and Hojjat Adeli, "Functional community analysis of brain: A new approach for EEG-based investigation of the brain pathology." *Neuroimage*, Vol. 58 (No. 2), pp. 401-08, (2011).

49- Richard Beare *et al.*, "Altered structural connectivity in ADHD: a network based analysis." *Brain imaging and behavior*, Vol. 11 (No. 3), pp. 846-58, (2017).

50- Teague R Henry and Jessica R Cohen, "Dysfunctional brain network organization in neurodevelopmental disorders." in *Connectomics*: *Elsevier*, (2019), pp. 83-100.

51- Amine Khadmaoui *et al.*, "MEG analysis of neural interactions in attention-deficit/hyperactivity disorder." *Computational intelligence and neuroscience*, Vol. 2016(2016).