Wavelet-Synchronization Methodology: A New Approach for EEG-Based Diagnosis of ADHD

Article *in* Clinical EEG and neuroscience: official journal of the EEG and Clinical Neuroscience Society (ENCS) · January 2010

DOI: 10.1177/155005941004100103 · Source: PubMed

2 authors, including:

Mehran Ahmadlou
Netherlands Institute for Neuroscience
29 PUBLICATIONS **943** CITATIONS

SEE PROFILE

Wavelet-Synchronization Methodology: A New Approach for EEG-Based Diagnosis of ADHD

Mehran Ahmadlou and Hojjat Adeli

Key Words

Attention-Deficit/Hyperactivity Disorder
Chaos Theory
Electroencephalography
Functional Connectivity
Generalized Synchronization
Radial Basis Function Neural Network
Synchronization Likelihood
Wavelet

ABSTRACT

A multi-paradigm methodology is presented for electroencephalogram (EEG) based diagnosis of Attention-Deficit/Hyperactivity Disorder (ADHD) through adroit integration of nonlinear science; wavelets, a signal processing technique; and neural networks, a pattern recognition technique. The selected nonlinear features are generalized synchronizations known as synchronization likelihoods (SL), both among all electrodes and among electrode pairs. The methodology consists of three parts: first detecting the more synchronized loci (group 1) and loci with more discriminative deficit connections (group 2). Using SLs among all electrodes, discriminative SLs in certain sub-bands are extracted. In part two, SLs are computed, not among all electrodes, but between loci of group 1 and loci of group 2 in all sub-bands and the band-limited EEG. This part leads to more accurate detection of deficit connections, and not just deficit areas, but more discriminative SLs in sub-bands with finer resolutions. In part three, a classification technique, radial basis function neural network, is used to distinguish ADHD from normal subjects. The methodology was applied to EEG data obtained from 47 ADHD and 7 control individuals with eyes closed. The Radial Basis Function (RBF) neural network classifier yielded a high accuracy of 95.6% for diagnosis of the ADHD in the feature space discovered in this research with a variance of 0.7%.

INTRODUCTION

Attention-Deficit/Hyperactivity Disorder (ADHD) is defined as a behavioral, cognitive and neuropsychological developmental disorder in childhood which can persist into adulthood.[1] ADHD is known as one of the most prevalent pediatric disorders.[2,3] Recent research indicates reduced dopamine efficacy in striatum, prefrontal and frontal areas, dysfunction in frontal-striatal-cerebellar circuitry, and impaired connectivity of frontal cortex as possible causes and sources of this disorder.[4-8] However, the causes and sources of ADHD have yet to be determined completely. With symptoms similar to many other deficits such as deficits in working memory, Learning Disorder (LD) and mood disorders the accurate diagnosis of ADHD becomes more complex.[9,10] Its main diagnostic criteria are: hyperactivity-impulsivity and inattentiveness. Impulsive action is defined as the inability to withhold an inappropriate response, as a deficient action in response inhibition, or as an action without foresight. A prevalent concept of inattentiveness is the inability to maintain or more generally regulate appropriate attention to a subject.[11] ADHD is classified into: predominantly inattentive, predominantly hyperactive-impulsive, and combined.[9] In this paper, we investigate ADHD as a comprehensive class irrespective of this classification.

ADHD-like behaviors are assessed primarily with visual discrimination tasks measuring the aforementioned characteristics. Also, attention, impulsivity, motivation, speed of responding, and several other criteria are tested and quantified by cognitive and behavioral tasks for recognizing ADHD-like behaviors.[1,9,12-14] Since the behavioral tasks may be controlled voluntarily and unknown subjective and even environmental factors can interfere with them, the results of the aforementioned methods cannot be sufficiently reliable.

EEGs are commonly used for diagnosis of epilepsy.[15-25] Recent research aims to use EEGs for diagnosis of other neurological disorders such as Alzheimer's disease.[26-29] EEGs as demonstrators of brain states may be a fast and reliable tool to aid in the diagnosis of ADHD. The eyes-closed EEG is preferred for the ADHD diagnosis because it reduces external stimulus which would approximate what in systems theory is called zero-input response where the brain is considered as the system. For the purpose of distinguishing subjects with ADHD from those without ADHD, the data used in this investigation are the eyes-closed EEG data.

In recent years, research has been performed on EEG-based diagnosis of ADHD mostly with stimulus (event-related EEG)[30,31] and some in the rest state (eyes-open and eyes-closed).[32,33] Ponomarev et al.[30] attempted to differentiate adolescents with ADHD from those without it by measuring evoked EEG synchronization in a movement inhibition task (Go/No-Go task). They calculated a measure of synchronization (a ratio of signal power variations at each sampling time moment to average power in the pre-stimulus interval) in 3 sub-bands from 4 Hz to 20 Hz which are known to be related to movement inhibition. For the selected feature they report p-values in the range of 0.02-0.03 for distinguishing between ADHD and non-ADHD subjects.

Using phase-locking analysis, Yordanov et al.[34] investigated relative differences of attention between ADHD and non-ADHD groups in a selective attentional task in certain band (gamma, 31-63 Hz) and in response to auditory stimuli. Kovatchev et al.[35] used a 3-dimensional power spectrum of EEG, and Fast Fourier Transformation (FFT), to

Mehran Ahmadlou is a Graduate Student from Amirkabir University of Technology, Tehran, Iran. Hojjat Adeli is Abba G. Lichtenstein Professor from the Departments of Biomedical Engineering, Biomedical Informatics, Civil and Environmental Engineering and Geodetic Science, Electrical and Computer Engineering, Neurological Surgery, and Neuroscience, Ohio State University, Columbus, Ohio.

Address request for reprints to Prof. Hojjat Adeli, Ohio State University, 470 Hitchcock Hall, 2070 Neil Avenue, Columbus Ohio 43210, USA.

Email: adeli.1@osu.edu

Received: September 2, 2009; accepted: November 11, 2009.

show that ADHD patients have deficits switching from one attention-demanded task to another one. Using a logistic regression model, they reported an accuracy of 80% for classifying ADHD. Using quantitative EEG (a statistical method of brain function evaluation based on brain electrical activity mapping of data obtained from 70 ADHD patients and 70 control subjects with eyes closed and resting) and electro-dermal activity (a measure of changes in electrical conductance of skin), Hermens et al.[32] found differences in the theta band of males and females with ADHD, independent of subtype.

Alexander et al.[31] found out that ADHD subjects, in their tasks with auditory and visual stimuli, had low EEG power and activity at frequencies around 1 Hz. Koehler et al.[33] assessed the power density of EEG in adult ADHD patients in some sub-bands at eyes-closed rest state. They reported alpha and theta power densities are different in ADHD and non-ADHD groups and found some differences in the EEG power spectrum between adult and children with ADHD.[33]

In this paper, a new ADHD diagnostic methodology is presented following the general diagnostic wavelet-chaos methodology developed recently by Adeli and his associates[15,28,36,37] and the ideology that neuronal connectivity deficits are determining factors in ADHD.[38]

The wavelet-chaos methodology seeks to extract chaotic and nonlinear dynamic features in EEG and its sub-bands from appropriate channels to classify a subject as having a deficit. Specifically, this methodology shows in a statistically meaningful way, how differences among chaotic properties may occur in certain sub-bands. Adeli et al.[37] developed this methodology successfully for diagnosing epilepsy and for seizure detection. Preliminary results have also been reported for the diagnosis of Alzheimer's disease.[28] The methodology has not been applied for diagnosis of ADHD, which is the focus of the current research. For the purpose of EEG-based diagnosis of ADHD, in this research using Synchronization Likelihood (SL)[39] as the nonlinear feature and wavelet decomposition,[40,41] the authors attempt to detect deficient functional connectivity of neo-cortex network in ADHD in all EEG sub-bands and full-band EEG.

METHODOLOGY

A generalized synchronization called SL[39] is used in this research. SL is a global criterion which quantifies the nonlinear similarities and interdependencies among activities of different neuronal assemblies which are captured by the EEG. In this research it is postulated that the neuronal network connectivity deficits may discriminate ADHD patients in certain loci and sub-bands, which may not be detectable in the full-band EEG. First a wavelet decomposition is performed on each of a 19-channel EEG. Since the deficits may not be detected accurately from the full-band EEG, the EEG data are decomposed into 5 sub-bands, and feature extractions are applied to each sub-band and to the full-band EEG.

Next, is the computation of multivariate SLs which quantifies the similarity of each electrode's signal to all others. The loci with more deficient connections have lower multivariate SLs in the ADHD group than in healthy cases (they are called the Deficient Loci or DL, in this study). This may be revealed not in the full-band EEG, but in a certain sub-band of the EEG. The main purpose of this step is to select appropriate channels. In this step, the loci whose multivariate SLs are more discriminative are chosen by the statistical one-way analysis of variations (ANOVA) test. These loci most probably have more deficient interactions and connections with the loci whose multivariate SLs are the highest in both ADHD and healthy groups, which are the most effective and have the key roles in both groups (they are called the Effective Loci or EL, in this study). As such, both DL and EL are chosen to be investigated in more detail in the next step where the bivariate SL quantifies the synchrony between each electrode from DL and each electrode from EL. In this step, because of having fewer loci, the bivariate SLs can be computed in narrower bands in an attempt to classify the ADHD and control groups with higher accuracy. Indeed, if the extracted bivariate SLs are not significantly discriminative, this step can be applied to increasingly narrower frequency bands. Multivariate SLs help select the proper electrodes and thereby detect the connectivity deficits so that more discriminative features with a finer resolution can be possible. Thereafter, computation of the bivariate SLs involves only two electrodes rather than all electrodes as in the case of the multivariate SLs. The methodology is explained step by step as follows.

Step 1 Data Preprocessing and Wavelet Analysis

1a) An Infinite Impulse Response (IIR) low pass filter is applied to obtain a band-limited EEG in the 0-60 Hz range.

1b) Using a 4-level wavelet decomposition[37] the following sub-bands are obtained: Gamma (30-60 Hz), Beta (15-30 Hz), Alpha (8-15 Hz), Theta (4-8 Hz), and Delta (0-4 Hz). A block diagram of the data preprocessing and wavelet analysis is shown in Figure 1.

Step 2 Signal Reconstruction

Using Takens' theorem[42] the state space of each EEG signal and its sub-bands (individually) is reconstructed as a trajectory of *d*-dimensional points so that each point is represented as follows:

$$X_{k,n} = (x_{k,n}, x_{k,n+T}, \ldots, x_{k,n+(d-1)T}) \quad (1)$$

where $x_{k,n}$ is the *n*th point of the EEG record obtained from the *k*th electrode ($k = 1,2,\ldots, M$; in which M is the number of electrodes), T is the delay or lag time and d is dimension of the state space. Consequently, the number of points in each trajectory or the number of embedding vectors is $N - (d-1)T$; where N is the number of sampling points of the EEG. Since synchronization should be calculated among all channels, an equal number of sampling points and the same lag time and dimension are chosen for all EEG signals and their sub-bands. The dimension should be chosen such that the reconstructed trajectories in the state space are not folded (compressed in a binding region) and self-intersected. Based on previous studies d and T are chosen to be the same for signals of all channels in order to be able to compare their SLs.[43,44] The minimum embedding dimension of a signal is calculated using Takens' Theorem as described in Jiang and Adeli.[45] A supremum of the minimum embedding dimensions of the EEG obtained from various channels is used as the dimension of their state spaces (d). It is consequent from Takens' theorem that considering this dimension for all EEG data prevents the trajectories to be folded. Each signal is then reconstructed in the d dimensional space to represent its dynamic without considerable loss of information.

Step 3 Synchronization Likelihood Computations Among All Electrodes (Multivariate SL)

In a time series, a pattern may repeat itself closely in the signal of each channel, whether a few or many times.[46] When there is generalized synchronization between two signals, a pattern present in one signal tends to repeat itself in the other signal at the same times. In this step, the likelihood and similarity between the trajectory of each channel and trajectories of all other channels are computed.

3a) For each EEG the probability that a point $X_{k,m}$ is closer to a reference state of the trajectory, $X_{k,n}$, than a distance $\varepsilon_{k,n}$ is calculated where n represents the index of a reference state and m is the index of a different trajectory point. The aforementioned probability is computed for electrode k and point n according to[39]:

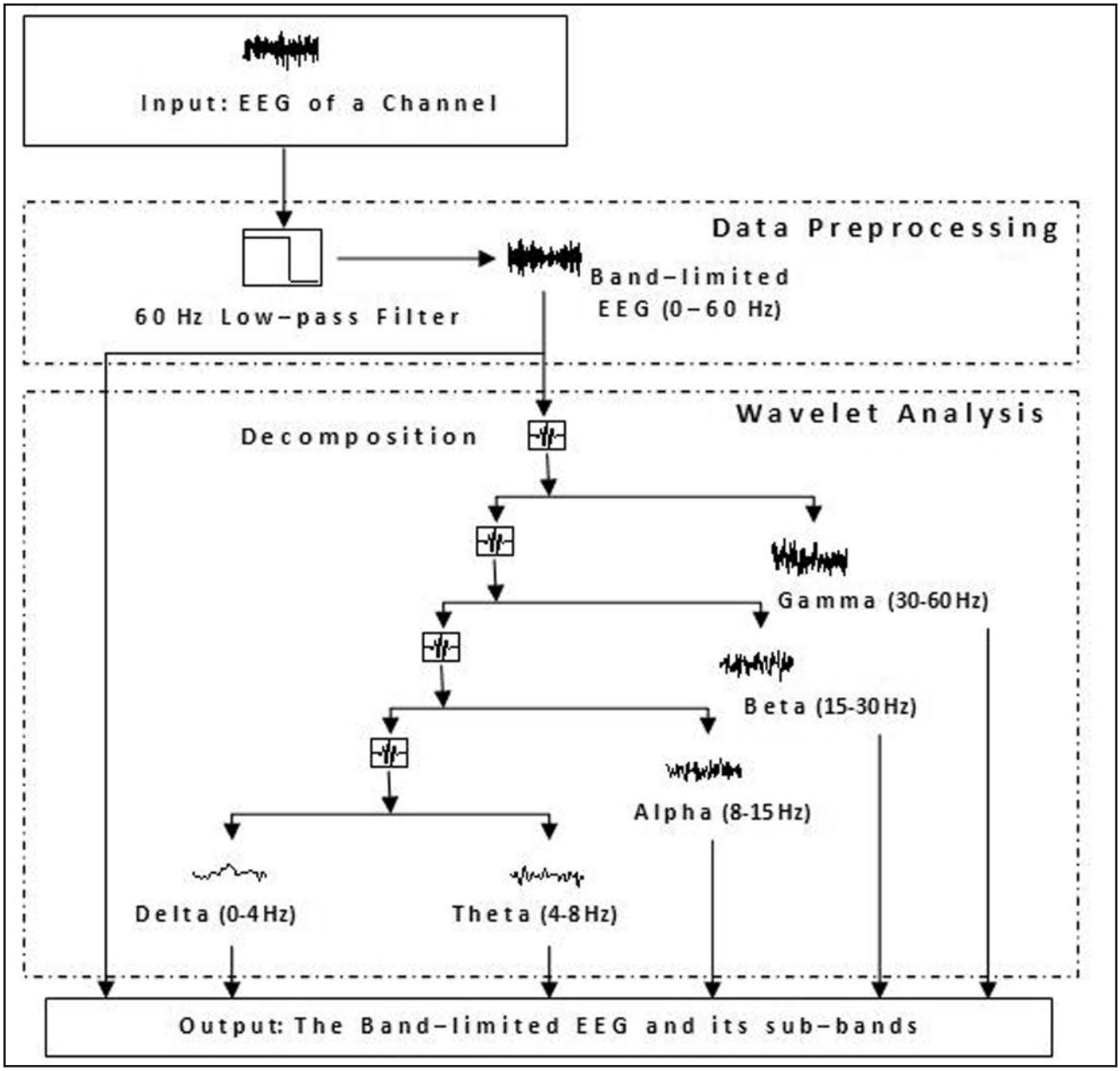


Figure 1.
Step 1 data preprocessing and wavelet analysis.

$$P_{k,n}^{\varepsilon_{k,n}} = \frac{1}{2(w_2 - w_1)} \sum_{\substack{m=1 \\ w_1 < |n-m| < w_2}}^{N} \Theta\left(\varepsilon_{k,n} - |X_{k,n} - X_{k,m}|\right) \qquad (2)$$

Where Θ is a Heaviside step function, and for all k, $\varepsilon_{k,n}$ is obtained by making $P_{k,n}^{\varepsilon_{k,n}}$ equal to a small P_{ref} ($P_{ref} << 1$). The sign | | in parentheses represents the norm. This probability function is used as a measure of similarity among channels. Also, this definition and measure makes the similarity computations independent of differences in the signals' amplitudes. The term $w_1$is the so-called Theiler correction.[42] Its purpose is to reduce the likelihhod of the vectors (states) representing a recurrence of the reference state ($X_{k,n}$) in order to avoid pseudo-similarities that may be not related to the dynamic system. Montez et al., (2006)[46] suggest a value of $(d - 1)T$ for w_1 to avoid overlap between the reference state and other states. The same value is used in this research. The parameter w_2 is related to the time interval or window (see Figure 2) in which the similarity of each state is compared with the reference state. For w_2, in this study a value equal to half of the sampling rate of the signal is used.

Then, $H_{n,m}$ (a positive integer) is calculated as the number of channels with Euclidean distance of their states, $X_{k,n}$ and $X_{k,m}$, less than distance $\varepsilon_{k,n}$:

$$H_{n,m} = \sum_{k=1}^{M} \Theta\left(\varepsilon_{k,n} - |X_{k,n} - X_{k,m}|\right) \qquad (3)$$

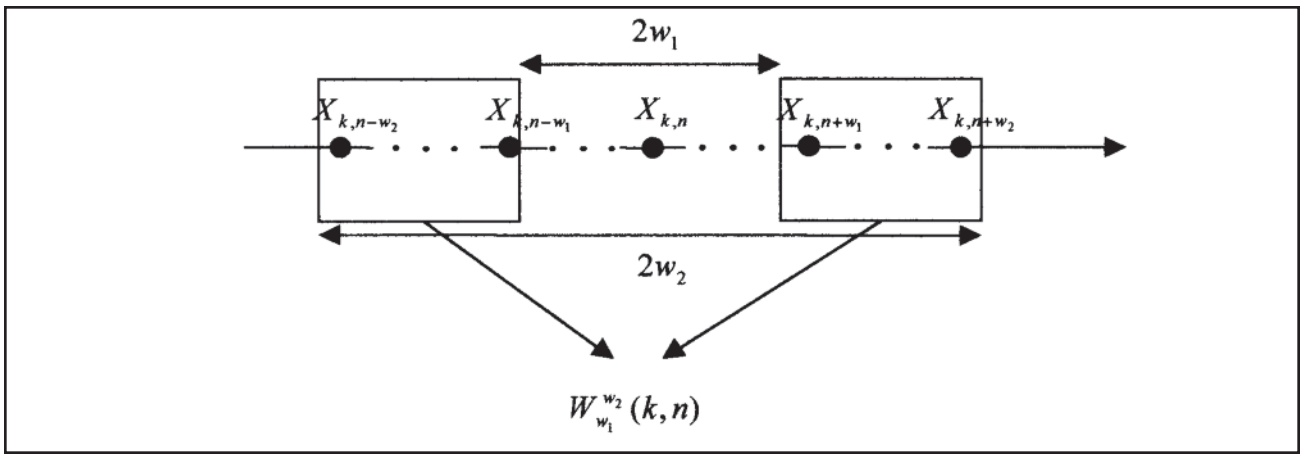


Figure 2.
Illustration of the window used in the SL computation, $X_{k,n}$ is the reference state and $2(w_2 - w_1)$ is the length of the window.

3b) Compute the synchronization likelihood for each channel bounded in the sliding window.[39]

$$S_{k,n} = \frac{1}{2P_{ref}(w_2 - w_1)} \sum_{\substack{m=1 \\ w_1<|n-m|<w_2}}^{N} S_{k,n,m} \quad (4)$$

$P_{ref} \leq S_{k,n} \leq 1$ determines how strongly channel k at time step n is synchronized with all other $M-1$ channels where the synchronization likelihood for each channel k and each discrete time pair (n,m) is defined as follows:

$$S_{k,n,m} = \frac{H_{n,m} - 1}{M - 1} \Theta(\varepsilon_{k,n} - |X_{k,n} - X_{k,m}|) \quad (5)$$

3c) The window is slid and the synchronization likelihood of the window is obtained. Consider q determines the distance the sliding window should go forward in each step. It generally can be considered any positive integer less than the number of points, $N - (d-1)T$, in the state space. In this research q is taken equal to a quarter of the sampling points of the signal. Therefore it means computing S_{k,n_0}, S_{k,n_0+q}, S_{k,n_0+2q},… (where n_0 is an initial counter equal to 1). Then, an average value of SL is computed for each channel (for the band-limited EEG and all its sub-bands, individually), which is the average of the calculated local similarities ($SL_k = mean\{S_{k,n_0}, S_{k,n_0+q}, S_{k,n_0+2q},\ldots\}$).

Step 4 Statistical Analysis

One-way ANOVA is employed for statistical analysis which determines how strongly a feature can discriminate groups based on variations both between and within groups. One-way ANOVA produces a number called p-value to show this ability of a feature. P-value is bounded between 0 and 1 where a low p-value indicates the ability of the feature to discriminate the groups and a high p-value shows similarity and closeness of distribution of the groups.

Using this statistical tool, the following channel and feature (multivariate SLs) selections are performed.

4a) A one-way ANOVA test is performed on computed SL of each band limited EEG and all its sub-bands obtained from each electrode to determine which features potentially have the ability to distinguish ADHD from control groups. For this purpose, the features for which the one-way ANOVA yields p-values less than a small number $0 < \varepsilon_1 < 0.1$ are considered. The SLs of the corresponding band limited EEGs or EEG sub-bands which satisfy this condition are chosen to construct the proper feature to be used in the next step of EEG classification. The discriminative SLs with *p-value* $\leq \varepsilon_1$ imply that the selected electrodes (which satisfy *p-value* $\leq \varepsilon_1$) reveal the loci that have the most deficient connections with the other loci. In this step the focus is on selection of deficient loci even though the sub-bands help detect the loci. More detailed information will be extracted from all sub-bands of the selected deficient loci in the next step (bivariate SL). The electrodes which satisfy the aforementioned statistical condition will be used to investigate the deficient connectivity in more details in step 5.

4b) A second parallel statistical analysis is performed between ADHD and control group to identify the most effective brain loci in both groups while in the resting state. In other words, the goal is to determine which channels/loci play the biggest role in brain's synchronizations in both groups. The high multivariate SL of a channel means that channel is strongly related to the states of other channels. The channels with high SLs indicate a key role in forming the state of the brain (which is the resting state in this research). Thus, in this step the channels with high SLs in both ADHD and control groups are selected. Areas with high SLs have strong plasticity and play a major role in the global functioning of the brain. Therefore deficient connections between the deficient areas obtained from step 4a with these areas play a major role in ADHD. For this purpose, the channels which satisfy the following two conditions are selected: 1) its SL is more than φ and 2) the p-value of the channel's SL for distinguishing the classes is more than λ; where $0 << \lambda < 1$ and φ is a little less than maximum of the SLs to select the most involved areas in synchronization in the rest state. The first condition results in detecting areas which have high SLs and in turn have high plasticity and major role in global synchronization of brain. The second condition results in detecting loci whose global synchronization are not significantly affected by ADHD. The parameters λ and φ are tuned manually so that at least one channel is obtained as the result of this step as a channel (or channels) which refers (refer) to healthy area with major role in global functioning of the brain in both ADHD and non-ADHD subjects. Figure 3 summarizes steps 2, 3 and 4 in one block diagram.

Since a fewer number of electrodes is chosen in this step (steps 4a and 4b), it is possible to calculate bivariate SLs in narrower bands (For example, 0-2 Hz; 2-4 Hz; 4-8Hz; 8-12Hz; … instead of just 5 physiologically known sub-bands 0-4Hz … and 30-60 Hz) in order to

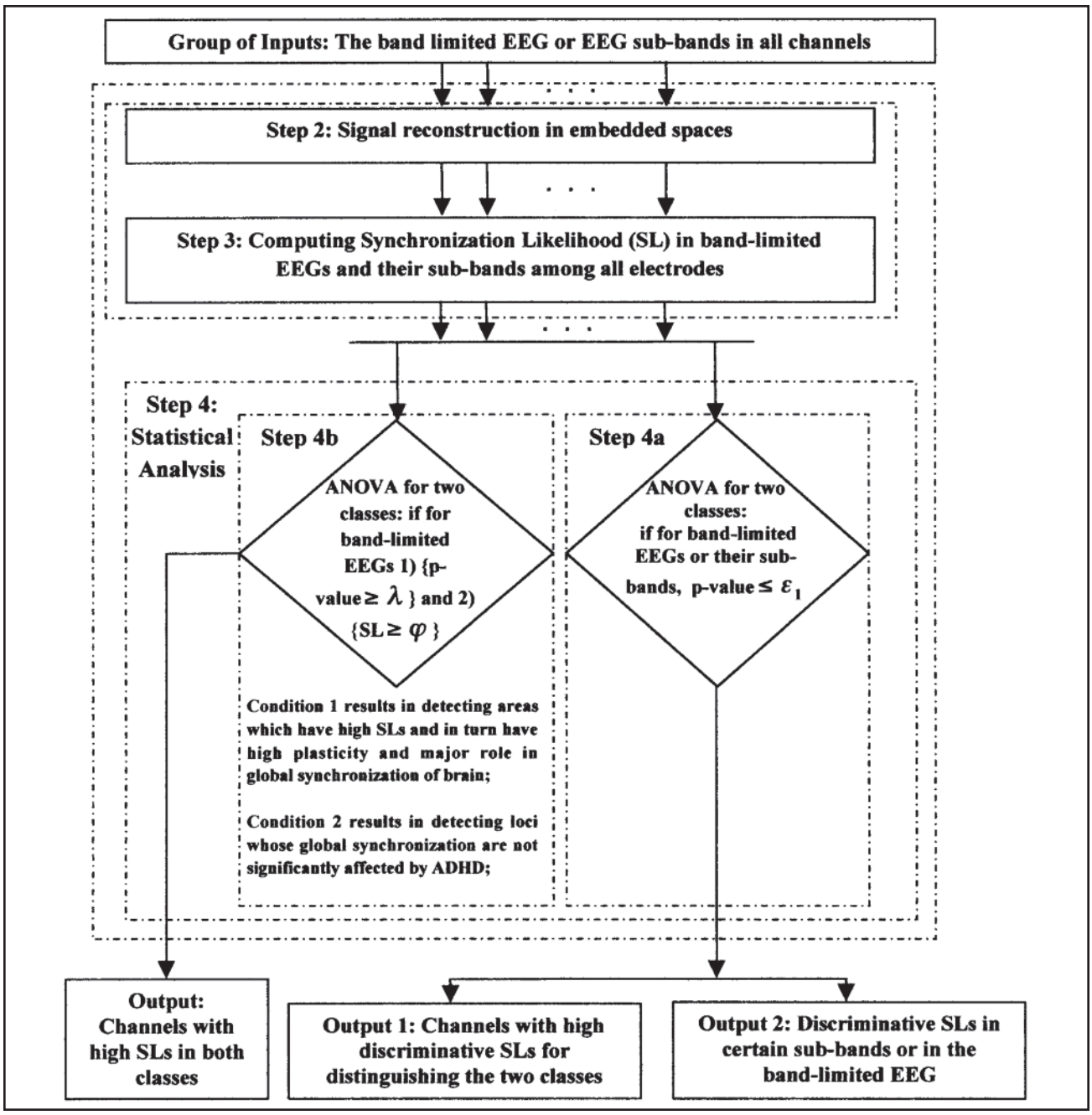


Figure 3.
Illustration of steps 2, 3, and 4 of the methodology: detecting primary locations of functional connectivity deficits and extracting primary features (SLs).

reach more detailed and more discriminative features. This general approach is not limited to the diagnosis of ADHD but may be applied to EEG-based diagnosis of other neurological disorders.

Step 5 Computation of SL for Each Channel Pairs (Bivariate SL)

To identify the deficit connections in ADHD patients which may be revealed in certain bands, SL between each pair of inputs is computed: input 1: the reconstructed EEG and all EEG sub-bands of the channels selected in Step 4a and input 2: the reconstructed EEG and all EEG sub-bands of the channels obtained in Step 4b. SL between each pair of trajectories is computed according to the following steps:

5a) Reconstruct the signals of each aforementioned pair of channels (channel k and channel r), in their embedding spaces, $X_{k,n}$, $X_{r,n}$, in the same manner as in Step 2. Then calculate $P^{\varepsilon}_{k,n}$ and $P^{\varepsilon}_{r,n}$, and ε^{n}_{k} and ε^{n}_{r} in the same manner as Step 3a.

5b) Calculate the synchronization between states of the trajectories of channels k and r, with reference states $X_{k,n}$ and $X_{r,n}$, respectively.

$$S_{(k,r),n} = \frac{1}{2P_{ref}(w_2 - w_1)} \sum_{\substack{m=1 \\ w_1<|n-m|<w_2}}^{N} \Theta(\varepsilon_{k,n} - |X_{k,n} - X_{k,m}|)\Theta(\varepsilon_{r,n} - |X_{r,n} - X_{r,m}|) \quad (6)$$

This equation is similar to Eq. (5) with M=2 which means two electrodes. Therefore *M-1=1* and $H_{n,m}=1$ (if the states $X_{k,m}$ and $X_{r,m}$ are not similar; which in turn means at least one of the Θs in Eq. 6 is equal to zero and therefore their product is equal to zero), or 2 (if the states $X_{k,m}$ and $X_{r,m}$ are similar; which in turn means both Θs in Eq. 6 are equal to one and their product is equal to one). In this case of *M=2* there are two channels k and r (between 1 and 19), but since $S_{k,n}=S_{r,n}$, the notation has been changed for both to $S_{(k,r),n}$ in a new form called bivariate SL between channels k and r in the related window (with center n).

Calculate $S_{(k,r),n_0}$, $S_{(k,r),n_0+q}$, $S_{(k,r),n_0+2q}$,… where *q* and n_0 are the same as step 3c. An average synchronization between channels *k* and *r* is computed by averaging the values obtained from Eq. (6) for all of the local similarities ($S_{(k,r)} = mean\{S_{(k,r),n_0}, S_{(k,r),n_0+q}, S_{(k,r),n_0+2q},\ldots\}$).

5c) Statistical Analysis: The one-way ANOVA is performed on all computed bivariate SLs between each pair of EEGs or their sub-bands to determine the ones with potentially high ability to discriminate the ADHD from control classes, that is, those which lead to *p-value*≤ ε_2 (considering a small value $\varepsilon_2 \le \varepsilon_1$). The reason for selecting a smaller value for ε_2 is the need for higher accuracy in detecting deficient connections.

Step 6 Classification

In this final step, EEG data are classified into control and ADHD groups. Both the discriminative multivariate SLs obtained from Step 4b and the discriminative bivariate SLs selected in Step 5, the more robust features, are used as the inputs of the classifier. The classifier used in this research is RBF neural network.[47-54]

The RBF neural network consists of three layers: an input layer, a hidden radial basis function layer, and an output layer, which has one neuron in this application. The number of nodes in the input layer is equal to the summation of the number of multivariate SLs discovered in Step 4a and the number of bivariate SLs of channel pairs discovered in Step 5. The number of nodes in the hidden layer is found based on the data used for training the network to yield the most accurate results as described in the next section. The output layer has one neuron with a hard limiter function which determines the data belongs to ADHD class or healthy class depending on the output is 1 or 0, respectively. The weighted input to the radial basis transfer functions of the hidden layer is the Euclidian distance between the input vector and the weight matrix of links connecting the input neurons to the hidden layer's nodes, multiplied by a bias value. As the distance between the weight vector and the input vector decreases, the output increases. The weights and biases are obtained by training the neural network to minimize the misclassification error in a Mean Square Error (MSE) sense.

Figure 4 presents the block diagram for steps 5 and 6.

DATA ACQUISITION

The data used in this research were recorded at the Atieh Comprehensive Center for Psych and Nerve Disorders, Tehran, Iran. The subjects signed a consent form. The EEG data were collected from 47 ADHD patients (7-12 years old, 7 girls; 40 boys) and 7 control individuals, (7-12 years old; 1 girl; 6 boys), (p-value of ages of the two groups = 0.98), with eyes closed and resting (sitting comfortably in a silent room with the same controlled temperature for all subjects). The ADHD group includes all clinical subtypes (hyperactive-impulsive, inattentive, and combined).

Scalp electrodes were applied according to the 10-20 system and all data were recorded with a sampling rate of 256 Hz. On the one hand, accurate computation of nonlinear features used in this study needs many sequential sample times. On the other hand, ADHD children cannot sit calm for a long time. Therefore, the EEG recording lasted 3 minutes (3x60x256=46,080 time samples) in order to have a large interval free from artifacts. An interval of 14,000 points of each EEG data was selected as an appropriate part of the recorded EEG without eye blink and other artifacts. Data for each patient or control subject consists of 19 channels and as such a total of 19x(47+7) = 1,026 EEGs was obtained from 54 subjects and used in this research.

APPLICATION AND RESULTS

Coifman wavelet of order 5 was used for wavelet decomposition in 4 levels to decompose the EEG to 5 sub-bands mentioned earlier in the paper (Figure 1). An embedding dimension of *d = 24* was used for all EEGs and EEG sub-bands for all channels based on the maximum of the embedding dimensions found for all channels and EEG data. Also, after trying several different values from 1 to 10 (recommended values in the literature), a time delay of *T=1* was used because it represented the higher ability of the extracted SLs in discriminating ADHD and non-ADHD EEGs at the final step.

In Step 2, all EEG sub-bands and the band-limited EEGs were reconstructed in their embedded space. In Step 3 values of w_2=128 and P_{ref}=0.05 were selected after trying several different values following the ranges of recommended values in the literature. The number of SLs calculated in Step 3 is equal to (47+7)x19x6 = 6,156 {(number of patients) x (number of channels) x (number of sub-bands plus one for the full-band EEG)}. The number of features, SLs, is 19x6 = 114 for each individual EEG record (with or without ADHD).

In Step 4a, a small number of loci with deficient connections are sought in order a) to investigate the connectivity deficits with higher resolutions and b) to minimize the high computational cost involved in the subsequent more detailed analysis. This would also lead to a fewer number of discriminative channels for EEG biofeedback application, the other objective of this research. In order to obtain a sufficiently small number of acceptably discriminative electrodes based on their multivariate SLs obtained in Step 4a, a value of ε_1=0.012 was obtained by trial and error. The one-way ANOVA showed the SL of certain channels in certain sub-bands can separate the 47 ADHD patient data from the 7 healthy control data under the described condition (p-value≤0.012). Table 1 shows the discriminating channels and sub-bands with p-value≤0.012 obtained in step 4a. Figure 5 demonstrates their SL values with their variations. Table 1 and Figure 5 show that the high distinctive de-synchronization occurs in O2, P4 and T5 loci of ADHD patients' brain in certain EEG sub-bands which in turn show connectivity deficits in those areas. Loci O2, T5, and P4 correspond respectively, to the right occipital, left temporal, and right parietal areas (Figure 6). The three areas are associated with the visual, and auditory, and data fusion centers, respectively. Detecting deficient connections between these areas and frontal areas in the rest state indicates deficits in neural plasticity, and neural paths between the mentioned areas. Therefore, these deficits could potentially reduce transmission of information between those areas. On the other hand, occipital, temporal and parietal areas are responsible for visual, auditory processing and data fusion and frontal areas are responsible for high level data processing. Therefore, a weak plasticity is inferred and neural paths found in this research implicitly show that sensory information in ADHD patients is not preprocessed as well as in the control group.

In Step 4b, since the goal is to select the channels with high synchronization likelihood, a relatively high value was chosen for φ, that is, equal to 80 percent of the maximum of the SLs obtained from each band-limited EEG (a total of 19 signals). In order to determine which EEG channels have high synchronization likelihood in both

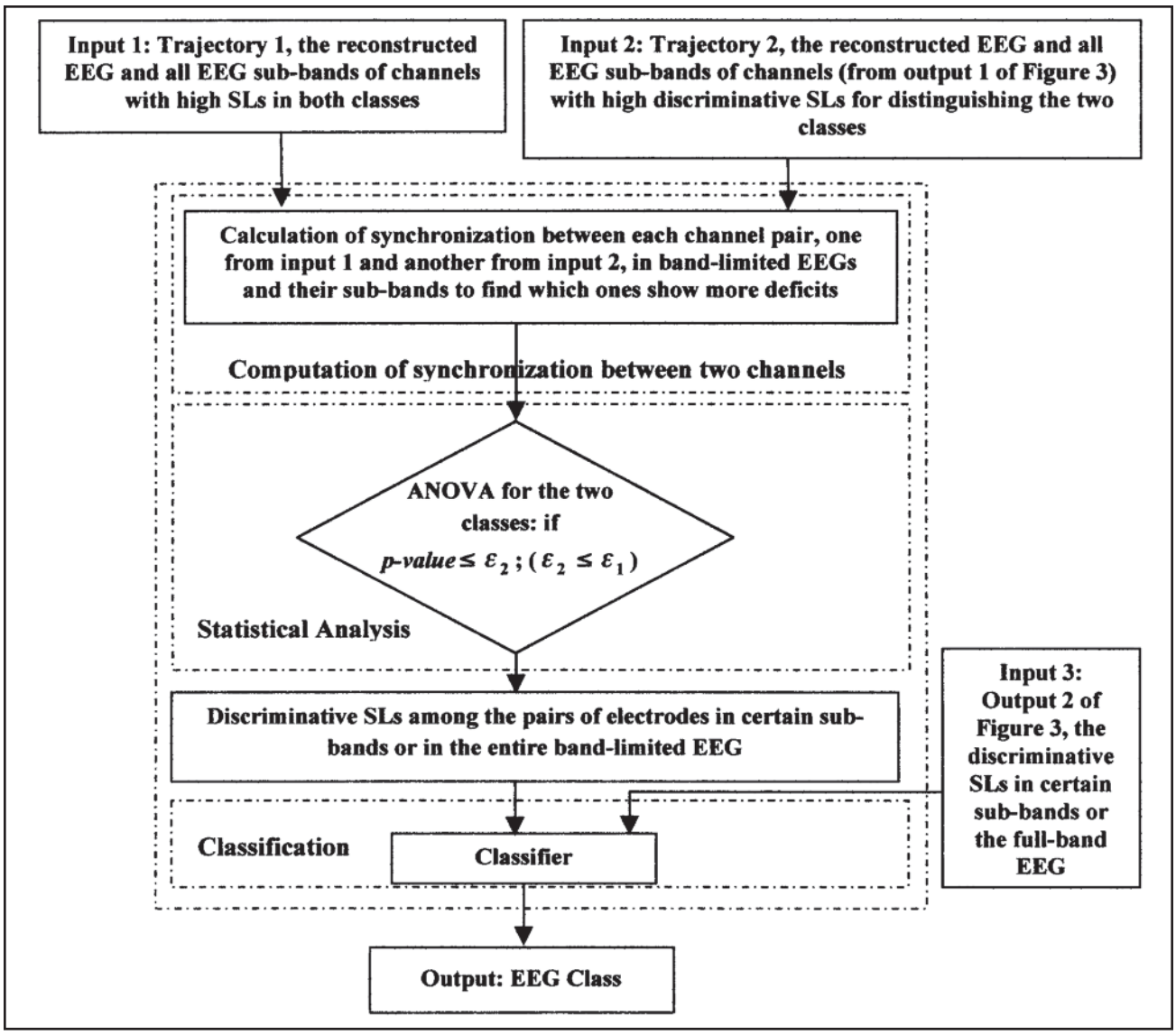


Figure 4.
Illustration of step 5 and 6: detecting locations of connectivity deficits and certain sub-bands as discriminative features to distinguish ADHD patients from control subjects.

ADHD and control groups (the most synchronous channels), a value of $\lambda = 0.8$ was chosen by trial and error. All channels detected by this procedure, a total of seven, are located in right and left 6 prefrontal and frontal lobes: Fp1, Fp2, F3, F4, F7, F8, and Fz as shown by gray circles in Figure 6. In general these areas are involved in pre-processing data and high level data processing in the brain.

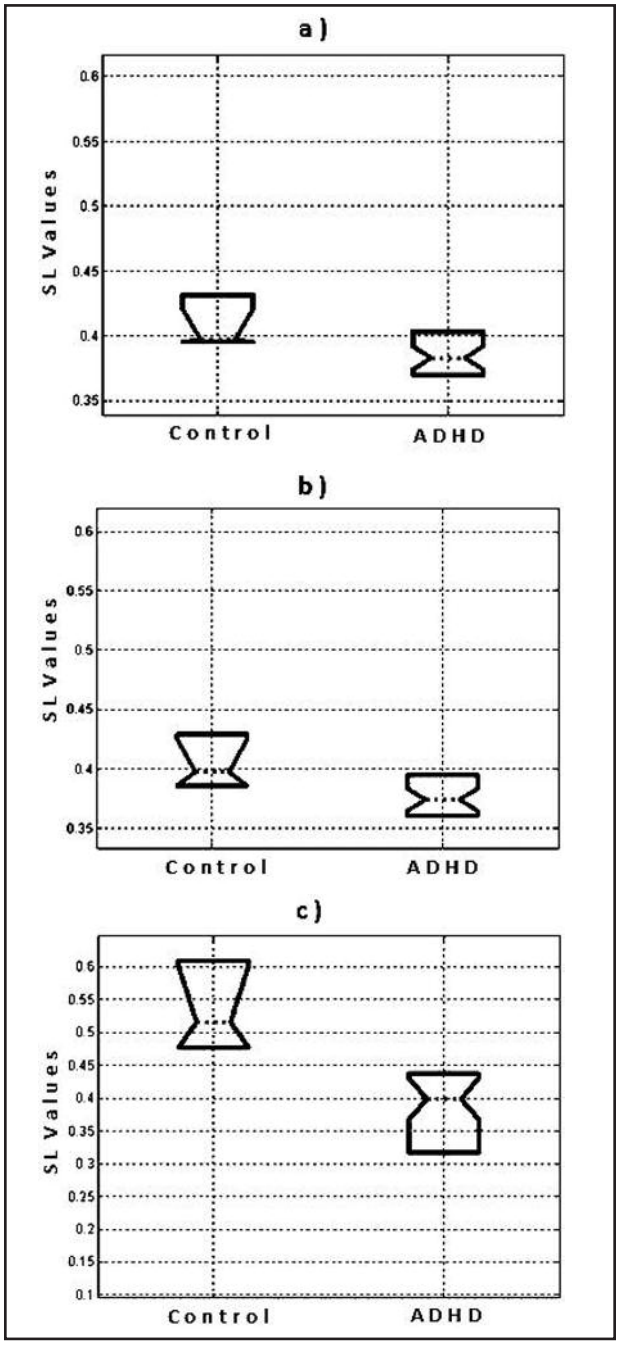
The most important finding in this research is that the detected prefrontal and frontal lobes have deficits in their connections with the associated sensory centers, O2, T5, and P4, (the connections are from the 3 channels, O2, T5, and P4, to the 7 channels, Fp1, Fp2, F3, F4, F7, F8, and Fz, that is, 3x7=21 connections). In Step 5, (21 connections) (5 sub-bands plus one for the full-band EEG) = 126 bivariate SLs were calculated among each pair of channels in each sub-band or in the entire band-limited EEG. The results of ANOVA test with p-value≤0.004 obtained in Step 5 are presented in Table 2.

Thirteen (13) features discovered in this research, the 3 multivariate SLs from Step 4a (presented in Table 1) and the 10 bivariate SLs of channel pairs (presented in Table 2) are used as the features of each individual's EEG data and the input vector of classifiers. As such, the Radial Basis Function (RBF) classifier has 13 input nodes. The data used in this research (based on the data described in the previous section) consists of 54 (47 ADHD patient data + 7 healthy control data) sets of 13-dimensional data. Ninety (90) percent of the data were selected randomly and used for training and the remaining 10 percent (6 data) were used for testing. In order to achieve a more reliable classification result, the test data includes the same number of ADHD patients and healthy control data (i.e., 3 ADHD patient data and 3 healthy control data). The classification results reported in this paper are based on repeating this random constrained selection 100 times and averaging the values in order to obtain statistically meaningful results.

Table 1

The channels and sub-bands for distinguishing ADHD from the control group with p≤0.012 obtained in step 4a

Channel	Sub-band	P-Value
02	Theta	0.0083
P4	Theta	0.012
T5	Delta	0.0103



Figures 5a-c.
The results obtained in step 4a: synchronication likelihood ranges in ADHD and control groups extracted from: a) theta band of channel O2, b) theta band of channel P4, c) delta band of channel T5.

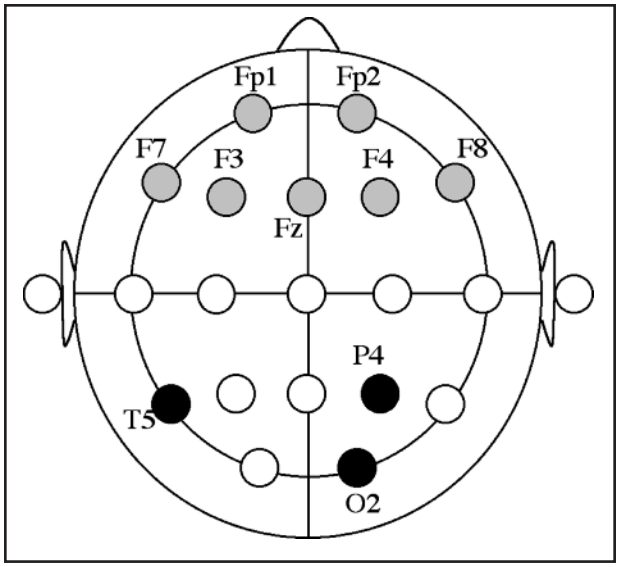


Figure 6.
Light gray circles indicate the loci with high SL and black circles indicate the loci that have deficits in their connections with other areas in ADHD patients.

As noted earlier, the classification has been done based on 13 features related to the connectivity assessment in the neo-cortex. The number of nodes in the hidden layer of the RBF classifier is 17. This number of nodes in the hidden layer produced accurate results (a smaller number would result in less accurate results and a larger number would not improve the accuracy any further). A spread parameter (which tunes the spread of the radial basis neurons) of 4.2 was found to yield sufficient convergence in minimizing the training and testing errors. The RBF neural network classifier yielded a high accuracy of 95.6% for diagnosis of the ADHD in the feature space discovered in this research with a variance of 0.7%.

CONCLUSIONS

Attempts on EEG-based diagnosis of ADHD so far have been based on primarily linear concepts such as power of EEG sub-bands, cross correlation, and coherence. These are not suitable for providing reliable and accurate results because EEG is a complex and chaotic signal. In this article, a multi-paradigm methodology was presented for EEG-based diagnosis of ADHD through adroit integration of nonlinear science, wavelets, a signal processing technique, and neural networks, a pattern recognition technique. The selected nonlinear features are generalized synchronizations known as synchronization likelihoods (SL), both among all electrodes and among electrode pairs. These features can detect the interdependencies among the spatial areas of neo-cortex at two level, the entire brain and details of frequency sub-bands. They can recognize the connectivity deficits in certain areas of the cortex.

The following observations can be gleaned from the literature of investigating coherency and synchrony of ADHD EEGs: 1) the published studies are mostly on event-related EEGs (with auditory and/or visual stimuli),[34,38] 2) researchers have used linear measurements of synchronization and coherence whether in certain sub-bands or in the full-band EEG,[30] 3) the coherency and synchrony have been calculated locally just between each pair of electrodes, and not among all electrodes as a synchrony of each electrode with all

Table 2

The impaired functional connections of pairs of channels in certain sub-bands obtained in step 5

Connection of pair of channels	Fp2, P4	F3, P4	Fz, O2	F4, T5	F4, T5	Fp1, P4	F3, P4	F8, P4	Fp2, O2	F7, O2
EEG sub-bands	Theta	Theta	Theta	Theta	Delta	Delta	Delta	Delta	Delta	Delta
p-value	0.0038	0.0027	0.0025	0.0024	0.0019	0.0040	0.0022	0.0022	0.0033	0.0008

others. To this end, some researchers have reported abnormal gamma synchrony in event-related EEGs of ADHD,[34] and some others have recognized abnormal lower and upper alpha synchrony in event-related EEG of ADHD.[38]

To the best of the authors' knowledge this is the first time chaos theory and nonlinear science has been used for diagnosis of ADHD. Another novelty of the proposed methodology, is that nonlinear features are detected not only from the full-band EEG, but also from EEG sub-bands. It is demonstrated that the discriminative features may not be detected in the full-band EEG, but may be recognized in certain sub-bands. In fact, the new methodology is a spatio-frequency technique aimed to detect both the deficit physiological connections in neo-cortex and discriminate nonlinear features which can be used to diagnose the ADHD based on eyes-closed EEG.

The new methodology consists of three main parts: first detecting the more synchronized loci (group 1: Fp1, Fp2, F3, F4, F7, F8, and Fz) and loci with more discriminative deficit connections (group 2: O2, P4, and T5). In this part, using SLs among all electrodes, discriminative SLs in certain sub-bands were extracted (SL of O2 in theta, SL of P4 in theta, and SL of T5 in delta). In part two, SLs are computed, not among all electrodes, but between loci of group 1 and loci of group 2 in all sub-bands and the full-band EEG. This part led to accurate detection of deficit connections, and not just deficit areas, and more discriminative SLs in sub-bands with finer resolutions. Since sufficiently accurate features (SLs) were discovered in the 5 sub-bands, there was no need to resolve the EEG into additional sub-bands. The one-way ANOVA was used to show the ability of the detected features to discriminate ADHD and healthy groups based on the EEG. In part three a classification technique is used to distinguish ADHD from normal subjects.

Another significant objective of the EEG-based diagnosis of ADHD in this research is extracting appropriate and robust features and discovering markers that may be used in treatments using EEG biofeedback to improve attention and reduce impulsivity. EEG biofeedback is currently used as an effective method of ADHD treatment. But up to now, all protocols of this method use linear features such as alpha to theta power and beta to theta power ratios. The existing approach ignores the complexity and nonlinearity of the neuronal activities and their interactions. The nonlinear discriminating features obtained from this study can be used to improve this therapeutic methodology. Since in EEG biofeedback, from both computational time and expense perspectives, reducing the number of electrodes with the most discriminating features is desirable, this research is also aimed towards this goal. That is, the feature space discovered in this research can be used to train and treat ADHD patients by the EEG-biofeedback method more reliably and effectively instead of the less accurate features currently in use.

DISCLOSURE AND CONFLICT OF INTEREST

M. Ahmadlou and H. Adeli have no conflicts of interest in relation to this article.

REFERENCES

1. Kuntsi J, Andreou P, Ma J, Börger NA. Testing assumptions for endophenotype studies in ADHD: reliability and validity of tasks in a general population sample. BMC Psychiatry 2005; 5:40.
2. Hoerger ML, Mace FC. A computerized test of self-control predicts classroom behavior. Journal of Applied Behavior Analysis 2006; 39: 147-59.
3. Williams J, Taylor E. The evolution of hyperactivity, impulsivity and cognitive diversity. Journal of the Royal Society Interface 2006; 3: 399-413.
4. Nigg J, Willcutt E, Doyle A, Sonuga-Barke E. Causal heterogeneity in ADHD: should there be neuropsychological subtypes? Biological Psychiatry 2005; 57: 1224-1230.
5. Frank MJ, Santamaria A, O'Reilly RC, Willcutt E. Testing computational models of Dopamine and Noradrenaline dysfunction in Attention Deficit/Hyperactivity Disorder. Neuropsychopharmacology 2007; 32: 1583-1599.
6. Gilbert DL, Wang Z, Sallee FR, Ridel KR, Merhar S, Zhang J, et al. Dopamine transporter genotype influences the physiological response to medication in ADHD. Brain 2006; 129: 2038-2046.
7. Plessen KJ, Bansal R, Zhu H, Whiteman R, Amat J, Quackenbush GA, et al. Hippocampus and amygdala morphology in Attention-Deficit/Hyperactivity Disorder. Archives of General Psychiatry 2006; 63: 795-807.
8. Danker JF, Duong TQ. Quantitative regional cerebral blood flow MRI of animal model of Attention-Deficit/Hyperactivity Disorder. Brain Research 2007; 1150: 217-224.
9. Solanto MV, Gilbert SN, Raj A, Zhu J, Pope-Boyd S, Stepak B, et al. Neurocognitive functioning in AD/HD, predominantly inattentive and combined subtypes. Journal of Abnormal Child Psychology 2007; 35: 729-744.
10. Mayes SD, Calhoun SL, Bixler EO, Vgontzas AN, Mahr F, Hillwig-Garcia J, et al. ADHD subtypes and comorbid anxiety, depression, and oppositional-defiant disorder: differences in sleep problems. Journal of Pediatric Psychology 2008; 34: 328-337.
11. Sagvolden T. Behavioral validation of the spontaneously hypertensive rat (SHR) as an animal model of attention-deficit/hyperactivity disorder (AD/HD). Neuroscience and Biobehavioral Reviews 2000; 24: 31-39.
12. Siklos S, Kerns KA. Assessing multitasking in children with ADHD using a modified six elements test. Archives of Clinical Neuropsychology 2004; 19: 347-361.
13. Winstanley CA, Eagle DM, Robbins TW. Behavioral models of impulsivity in relation to ADHD: translation between clinical and preclinical studies. Clinical Psychology Review 2006; 26: 379-395.

14. Bitsakou P, Psychogiou L, Thompson M, Sonuga-Barke EJS. Inhibitory deficits in attention-deficit/hyperactivity disorder are independent of basic processing efficiency and IQ. Journal of Neural Transmission 2008; 115: 261-268.
15. Ghosh-Dastidar S, Adeli H, Dadmehr N. Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection. IEEE Transactions on Biomedical Engineering 2007; 54: 1545-1551.
16. Lee H, Cichocki A, Choi S. Nonnegative matrix factorization for motor imagery EEG classification. International Journal of Neural Systems 2007; 17: 305-317.
17. Osterhage H, Mormann F, Wagner T, Lehnertz K. Measuring the directionality of coupling: phase versus state space dynamics and application to EEG time series. International Journal of Neural Systems 2007; 17: 139-148.
18. Ghosh-Dastidar S, Adeli H. Improved spiking neural networks for EEG classification and epilepsy and seizure detection. Integrated Computer-Aided Engineering 2007; 14: 187-212.
19. Good LB, Sabesan S, Marsh ST, Tsakalis KS, Iasemidis LD. Control of synchronization of brain dynamics leads to control of epileptic seizures in rodents. International Journal of Neural Systems 2009; 19: 173-196.
20. Hamani C, Andrade D, Hodaie M, Wennberg R, Lozano A. Deep brain stimulation for the treatment of epilepsy. International Journal of Neural Systems 2009; 19: 213-226.
21. Jedlicka SS, Dadarlat M, Hassell T, Lin Y, Young A, Zhang M, et al. Calibration of neurotransmitter release from neural cells for therapeutic implants. International Journal of Neural Systems 2009; 19: 197-212.
22. Osorio I, Frei MG. Seizure abatement with single DC pulses: is phase resetting at play? International Journal of Neural Systems 2009; 19: 149-156.
23. Pineau J, Guez A, Vincent R, Avo M. Treating epilepsy via adaptive neurostimulation: a reinforcement learning approach. International Journal of Neural Systems 2009; 19: 227-240.
24. Shoeb A, Guttag J, Pang T, Schachter S. Non-invasive computerized system for automatically initiating vagus nerve stimulation following patient-specific detection of seizures or epileptiform discharges. International Journal of Neural Systems 2009; 19: 157-172.
25. Velasco AL, Velasco F, Velasco M, Nunez JM, David Trejo, García I. Neuromodulation of epileptic foci in patients with non-lesional refractory motor epilepsy. International Journal of Neural Systems 2009; 19: 139-147.
26. Adeli H, Ghosh-Dastidar S, Dadmehr N. Alzheimer's disease and models of computation: imaging, classification, and neural models. Journal of Alzheimer's Disease 2005; 7: 187-199.
27. Adeli H, Ghosh-Dastidar S, Dadmehr N. Alzheimer's disease: models of computation and analysis of EEGs. Clinical EEG and Neuroscience 2005; 36: 131-140.
28. Adeli H, Ghosh-Dastidar S, Dadmehr N. A spatio-temporal wavelet-chaos methodology for EEG-based diagnosis of Alzheimer's disease. Neuroscience Letters 2008; 444: 190-194.
29. Kramer MA, Chang FL, Cohen ME, Hudson D, Szeri AJ. Synchronization measures of the scalp EEG can discriminate healthy from Alzheimer's subjects. International Journal of Neural Systems, 2007; 17: 61-69.
30. Ponomarev VA, Kropotova OV, Kropotov YD, Polyakov YI. Desynchronization and synchronization of EEG related to stimuli triggering or forbidding sensory-motor reaction in adolescents: II. The peculiarities in case of the Attention Deficit and Hyperactivity syndrome. Human Physiology 2000; 26: 251-257.
31. Alexander DM, Hermens DF, Keage HAD, Clark CR, Williams LM, Kohn MR, et al. Event-related wave activity in the EEG provides new marker of ADHD. Clinical Neurophysiology 2007; 119: 163-179.
32. Hermens DF, Kohn MR, Clarke SD, Gordon E, Williams LM. Sex differences in adolescent ADHD: findings from concurrent EEG and EDA. Clinical Neurophysiology 2005; 116: 1455-1463.
33. Koehler S, Lauer P, Schreppel T, Jacob C, Heine M, Boreatti-Hümmer A, et al. Increased EEG power density in alpha and theta bands in adult ADHD patients. Journal of Neural Transmission 2009; 116: 97-104.
34. Yordanov J, Banaschewski T, Kolev V, Woerner W, Rothenberger A. Abnormal early stages of task stimulus processing in children with attention-deficit hyperactivity disorder: evidence from event-related gamma oscillations. Clinical Neurophysiology 2001; 112: 1119-1127.
35. Kovatchev B, Cox D, Hill R, Reeve R, Robeva R, Loboschefski T. A psychophysiological marker of Attention Deficit/Hyperactivity Disorder (ADHD)-defining the EEG consistency index. Applied Psychophysiology and Biofeedback 2001; 26: 127-140.
36. Adeli H, Zhou Z, Dadmehr N. Analysis of EEG records in an epileptic patient using wavelet transform. Journal of Neuroscience Methods 2003; 123: 69-87.
37. Adeli H, Ghosh-Dastidar S, Dadmehr N. A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy. IEEE Transactions on Biomedical Engineering 2007; 54: 205-211.
38. Murias M, Swanson JM, Srinivasan R. Functional connectivity of frontal cortex in healthy and ADHD children reflected in EEG coherence. Cerebral Cortex 2007; 17: 1788-1799.
39. Stam CJ, van Dijk, BW. Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets. Physica D: Nonlinear Phenomena 2002; 163: 236-251.
40. Zou W, Chi Z, Lo KC. Improvement of image classification using wavelet coefficients with structured-based neural network. International Journal of Neural Systems 2008; 18: 195-205.
41. Rizzi M, D'Aloia M, Castagnolo B. Computer aided detection of microcalcifications in digital mammograms adopting a wavelet decomposition. Integrated Computer-Aided Engineering 2009; 16: 91-103.
42. Chan K-S, Tong H. Chaos: A Statistical Perspective. New York: Springer; 2001: 113-122.
43. Arnold J, Grassberger P, Lehnertz K, Elger CE. A robust method for detecting interdependences: application to intracranially recorded EEG. Physica D 1999; 134: 419-430.
44. Pereda E, Rial R, Gamundi A, González J. Assessment of changing interdependencies between human electroencephalograms using nonlinear methods. Physica D 2001; 148: 147-158.
45. Jiang X, Adeli H. Fuzzy clustering approach for accurate embedding dimension identification in chaotic time series. Integrated Computer-Aided Engineering 2003; 10: 287-302.
46. Montez T, Linkenkaer-Hansen K, van Dijk BW, Stam CJ. Synchronization likelihood with explicit time-frequency priors. NeuroImage 2006; 33: 1117-1125.
47. Adeli H, Wu M. Regularization neural network for construction cost estimation. Journal of Construction Engineering and Management 1998; 124: 18-24.
48. Adeli H, Karim A. Fuzzy-wavelet RBFNN model for freeway incident detection. Journal of Transportation Engineering 2000; 126:464-471.
49. Karim A, Adeli H. Comparison of the fuzzy-wavelet RBFNN freeway incident detection model with the California algorithm. Journal of Transportation Engineering 2002; 128: 21-30.
50. Karim A, Adeli H. Radial basis function neural network for work zone capacity and queue estimation. Journal of Transportation Engineering 2003; 129: 494-503.
51. Ghosh-Dastidar S, Adeli H, Dadmehr N. Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection. IEEE Transactions on Biomedical Engineering 2008; 55: 512-51.
52. Pedrycz W, Rai R, Zurada J. Experience-Consistent Modeling for Radial Basis Function Neural Networks. International Journal of Neural Systems 2008; 18: 279-292.
53. Anand P, Siva Prasad BVN, Venkateswarlu C. Modeling and optimization of a pharmaceutical formulation system using radial basis function network. International Journal of Neural Systems 2009; 19: 127-136.
54. Savitha R, Suresh S, Sundararajan N. A fully complex-valued radial basis function network and its learning algorithm. International Journal of Neural Systems 2009; 19: 253-267.