Synchronization coupling investigation using ICA cluster analysis in resting MEG signals in Reading Difficulties

Marios Antonakakis, Giorgos Giannakakis, Manolis Tsiknakis, Sifis Micheloyannis and Michalis Zervakis

Abstract - **The understanding of the mechanisms of human brain is a demanding issue for neuroscience research. Physiological studies acknowledge the usefulness of synchronization coupling in the study of dysfunctions associated with reading difficulties. Magnetoencephalogram (MEG) is a useful tool towards this direction having been assessed for its superior accuracy over other modalities. In this paper we consider synchronization features for identifying brain operations. Independent Component Analysis (ICA) is applied on MEG surface signals in controls and children with reading difficulties and are clustered to representative components. Then, coupling measures of mutual information and partial directed coherence are estimated in order to reveal dysfunction of cerebral networks and its related coordination.**

I. Introduction

RESTING state electroencephalogram (EEG) has been used in many studies regarding neurological and/or psychiatric diseases. Synchronization measures and notions from graph theory are employed to characterize connectivity data from structural MRI, diffusion MRI, functional MRI, electroencephalogram and magnetoencephalogram (MEG) [1]. In particular, resting-state functional connectivity has been studied in relation with depression [2], Alzheimer disease (AD) [3], and schizophrenia [4]. Investigation of connectivity patterns during resting state (with eyes closed) between controls and dyslexic children using MEG may reveal valuable information, especially from the analysis of independent components which are less affected from brain conduction effects [5].

A well-established tool in extracting activities mainly in multichannel electroencephalogram (EEG) is Independent Component Analysis (ICA) [5], [6]. Via ICA a multivariate signal is separated into additive components, under the assumption of mutual statistical independence of non-Gaussian source signals.

Various synchronization features have been used in order to estimate coupling interactions between brain regions from neuromagnetic recordings of resting-state brain activity [7]. Bivariate and multivariate coupling measures such as mutual information (MI) [8] and partial directed coherence (PDC) [9] are used in this study. MI is a bivariate measure based on information-theory concepts that suffers from the problem of detecting artificial synchronization between conditionally independent variables. This problem is addressed by using multivariate measures [10].

The synchronization measures have been widely applied to EEG/MEG channels, convey artificial links of close-by regions due to brain conduction effect and the diffusion of energy in neighboring sites through the process of wave transmission in the brain conductance matter. Furthermore, they do not describe the actual brain activity, but just the reflection of such activity on the outskirts of the scalp. Thus, they cannot be related to actual sources in the brain. In the case of MEG, the existence of many channels induces quite complex networks with many nodes, whose examination and validation of interrelationships becomes quite difficult. This synchronization study involves estimates the independent components of the multichannel MEG, producing the main sources of activity and considering their direct interrelationships. Then, the components are clustered based on their topology and only one representative component per cluster is considered in the synchronization network, thus relieving the network structure from excess nodes. Overall, ICA is applied on MEG resting state recordings and is followed by a clustering approach grouping independent components (ICs). This aims to reduce dimensionality of initial data and organize clusters of ICA components according to their neurophysiological brain process stemming from different locations. In this form, the study of the MEG electrode signals can be efficiently used for the study of brain activity and the comparison of brain states or populations. In our specific analysis we perform a comparative analysis of synchronization measures of MI and PDC on clustered ICs, extracting and studying the characteristics of two groups, one control and the other of children with dyslexia.

II. Materials and Methods

A. *Participants and Preprocessing of MEG*

The study includes 26 children (12 males, 14 females, mean age 12.2±2.1 years) diagnosed with reading difficulties (RD group) and 40 children (25 males, 15 females, mean age 11±2.3 years) who had never experienced difficulties in

Manuscript received July 30, 2013.

M. Antonakakis, M. Zervakis are with the Technical University of Crete, (phone: +302821032716; e-mail: mantonakakis@isc.tuc.gr, michalis@display.tuc.gr).

G. Giannakakis, M. Tsiknakis, are with the Institute of Computer Science, Foundation for Research and Technology, Heraklion 71110, Greece (e-mail: {ggian; tsiknaki}@ics.forth.gr).

M. Tsiknakis is also with the Department of Applied Informatics and Multimedia, Technological Educational Institute of Crete.

S. Michelogiannis is with University of Crete, Faculty of Medicine, Greece (e-mail: micheloyannis.sifis@gmail.com)

978-1-4799-3163-7/13/$31.00 ©2013 IEEE

reading (NI group) served as controls group. Participants are categorized as having reading disabilities if they score below the 16th percentile of the standard score of 85 on the Basic Reading composite (average of Word Attack and Letter-Word Identification subtest scores of the Woodcock-Johnson Tests of Achievement-III [5]. All participants had Full Scale IQ scores above 80 on the Wechsler Abbreviated Scale of Intelligence [5] where participants were excluded if there was included history of neurological or psychiatric disorder (including ADHD). A more detailed data description can be found in [5], [11]. The study has been approved by the University of Texas, Institutional Review Board.

MEG recordings were obtained with a whole-head neuromagnetometer array (4-D Neuroimaging, Magnes WH3600), consisting of 248 first order axial gradiometer coils, which were in a magnetically shielded chamber. There were 3 minutes (16 subjects for NI group and 17 for RD group) and 5 minutes (24 subjects for NI group and 9 for RD group) of continuous MEG data were acquired at 1017.25 HZ. The preprocessing stage of raw MEG data included application of a notch filter for removing of power line nose (PLN) following by a low-pass filter at 58 Hz with an IIR Butterworth filter.

B. *Independent Component Analysis (ICA)*

Independent Component Analysis (ICA) reveals independent sources of activity from different MEG signals and attempts to separate the corresponding generators of MEG rhythms [12]. According to the basic model $S = W \cdot X$ (1) where **S** is the matrix with independent components (source table), **X** the matrix with mixture MEG signals and **W** called unmixing table (inverse of W results the mixing table **A**). The matrix **W** is calculated by means of Infomax algorithm, which is an iteration procedure that maximizes the mutual information between sources [12]. In order to compare characteristics for both groups on ICs the assumption described in [6], [13] was used regarding matrix W considering similar scalps for NI and RD group. The assumption imposes that cortical localization of components will be similar between the two groups. It implies that the unmixing matrix **W** will be common for group NI and RD, solving also the matching problem between scalp topographies. The unmixing matrix W was created only by using data from NI group as a gold standard for normal activity. Timeseries from this group were concatenated to a common matrix with 248 channels x 5127368 time points. Dimensionality reduction was first performed preserving only the most dominant components with the most significant information (above 95% of total energy for all subjects) according to Principal Component Analysis (PCA) [12]. The new dimensionality was 40 principal components (PCs) x 5127368 points. The ICA decomposition was performed using Fieldtrip toolbox [14] (Infomax algorithm [12]). The ICA finds the matrix **W** [40x248] and then it multiplied with every matrix from NI and RD group generating the source matrix **S**.

C. *Artifact detection-elimination and Clustering of ICs*

The application of ICA is a basic procedure for the recovery of cerebral and non-cerebral signals [12][15] but it is not clear which signals describe brain activity. A procedure for the classification of components such as ocular activity (OA), cardiac activity (CA), head moving and Gaussian noise components (GNC) is developed here, relying on the investigation of [15]. The procedure calculates the measures of kurtosis, Shannon entropy and skewness with a “segment two-step procedure considering approach” and a “global approach” [16]. The first step divides ICs to segments and calculates kurtosis, Shannon entropy and skewness. The second step calculates measures for the whole signal. The procedure detects ICs as OA, CA and GNC based on values of previous metrics and on the scalps topographies. If values from segment and global method for the metrics kurtosis and Shannon Entropy are over a threshold then signals recorded as OA, head moving and if values of skew over a threshold then signals recorded as CA. Thresholds are evaluated by Chebyshev’s inequality for every metric. If value of global kurtosis is very close to zero, IC detect as GNC.

One limitation associated with OA artifacts is that a straightforward removal of the ICs containing such activity most likely results in some loss of data, because those ICs rarely consist of blink-related EOG activity solely. To retain the brain activity in those selected ICs containing blinks, empirical mode decomposition (EMD) is used for the filtering of signals with OA [17]. In particular, we used EMD parabolic decomposition with extrapolated extrema - rEMD [18] because the classic algorithm of EMD suffered from problems [18].

The decision of artifacts and their removal is based on the combination of previous method and the clustering of ICs based on scalp topography (a similar but not same procedure at [19]). The row of matrix **A** is used for the representation of influence for every IC on scalp topography (based on fieldtrip toolbox [14]). In addition, matrix **A** is common for all dataset and it gives the advantage of common clustering result. In particular, we determined clusters using hierarchical clustering [20] with cosine similarity as a distance metric [21] applied on columns of **A**. The dendrogram used to illustrate the arrangement of the clusters produced by hierarchical clustering. A threshold defined from the percentage of the biggest distance between two existing clusters on dendrogram is used to decide the number of subtrees of dendrogram [20].

It has been reported that ICA can separate strong activity into different components but weaker activity is mixed within the rest of components [22]. This is particularly evident in EEG where the number of channels and components is very small. Even in case of MEG we observe a similar mixing of activity in multiple components. Notice that the number of components is reduced by the PCA process prior to the ICA application. After removing some of the components and return back to the electrode signals, the MEG is artifact free and a repetition of ICA allows more components to be decomposed as single activity signals. Thus, we expect that a second application of ICA allows for a better decomposition of brain activity into the available components. For this reason, we repeat the process and retain the new components for further consideration.

D. Application of synchronization measures on ICs

The identification of interdependence and information metrics is performed on the calculated IC components. First, we use a clustering approach for the concentration of ICs with similar scalp topography as shown in Fig. 1.

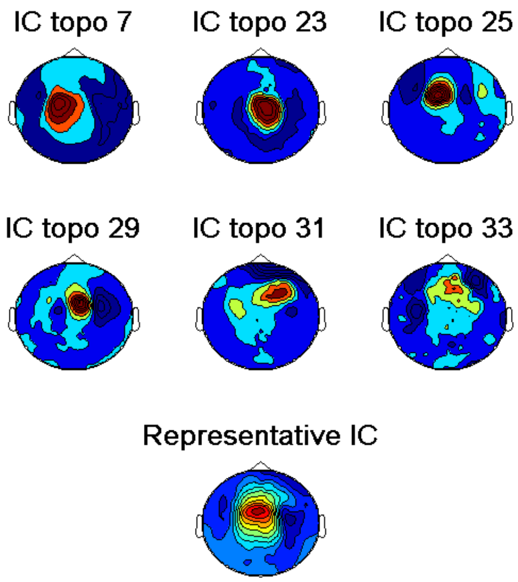

Fig. 1. A cluster of ICs and its representative IC.

One IC is preserved as representative for every cluster which calculated by means of localization and signals on every cluster [19]. Then the estimation of PDC and MI is performed on the representative ICs. MI is calculated on representative ICs separated with a band-pass Butterworth filter 5^{th} order on frequency bands: δ (0.5-4 Hz), θ (4-8 Hz), α_1 (8-10 Hz), α_2 (10-13 Hz), β (13-30 Hz) and γ (30-58 Hz). Alpha sub-bands were explored by the method proposed in [23].

E. Synchronization Coupling measures

A description of the synchronization measures used is provided below.

1) Mutual Information

Mutual information (MI) measures the interdependence of two time series using an algorithm derived from information-theory concepts. As measure of interdependence, MI provides many advantages based on [9]. The mathematical definition of MI, given by

$$I(X;Y) = \sum_{y\in Y}\sum_{x\in X} p(x,y)\log\left(\frac{p(x,y)}{p(x)p(y)}\right) \qquad (2)$$

with X and Y are two discrete random variables, $p(x,y)$ is the joint probability distribution function of X and Y and $p_x(x) = \sum_{y\in Y} p(x,y)$ and $p_y(y) = \sum_{x\in X} p(x,y)$ are the marginal probability distribution functions of X and Y, respectively.

2) Partial Directed Coherence

Directed information flow can be revealed by means of Partial Directed Coherence (PDC) which was introduced Baccala and Sameshima [10].

$$\pi_{i\leftarrow j} = \frac{\left|\bar{\mathrm{A}}_{\mathbf{ij}}\right|}{\sqrt{\sum_{l=1}^{n}\left|\bar{\mathrm{A}}_{\mathbf{lj}}(\omega)\right|^{2}}} \qquad (3)$$

where $\mathbf{A}$ is the $n\times n$ multivariate autoregressive matrix in frequency domain. The PDC $\pi_{i\leftarrow j}(\omega)$ takes values between 0 and 1 indicating directed coupling strength.

The model order p was estimated using Akaike information criterion and in all dataset it was a number between 20 and 22. Coefficients of matrix A were estimated by Partial Correlation Estimation from Nutall-Strand [24] with unbiased correlation function. Furthermore, it calculated integrated PDC – IPDC by

$$_{f_1}^{f_2}IPDC_{ij} = \int_{f_1}^{f_2}\pi_{ij}(f)df = \sum_{f_1}^{f_2}\frac{\left|\bar{\mathrm{A}}_{\mathbf{ij}}\right|}{\sqrt{\sum_{l=1}^{n}\left|\bar{\mathrm{A}}_{\mathbf{lj}}(\omega)\right|^{2}}} \qquad (4)$$

for every brain rhythm.

F. Statistical Analysis

Statistical analysis was performed on values of synchronization measures using SPSS. Distribution of IPDC values was normal but distribution of MI values was not normal.

As a result of normality test, two sample t-test used on IPDC values and the non-parametric Mann-Whitney test used for assessing statistical significance of differences of MI values. The selection of t-test with equal or no equal variances assumed estimated using F-test. The statistical analysis performed for every synchronization value of IC between NI and RD group per brain rhythm. The threshold for significant level of p-value defined at 95% level of confidence.

III. Results

A. Parameter selection

ICA aims to find a set of spatial filters that inverts the mixture of data and recovers the independent sources of brain activity from MEG signals. The mixing matrix, **A**, is estimated with a specific way [6], [13] having the advantage of common brain mapping on scalp for all dataset. Dimensionality reduction was used for the elimination of external noise [12]. The number of PCs was 40 which decided using a combination of scree plot and total variance explained by eigenvalues which is above 95% of total energy for all subjects.

An IC is detected as artifact, if 30% of values of metrics are over the thresholds according to the ‘segment approach’ and the ‘global approach’ [16]. Threshold decided using standard deviation (SD) according to Chebyshev’s inequality. Non-cerebral IMFs, from spatial decomposition of OA, are detected and removed using a threshold which calculated according to the SD of first IMF [17].

The decision of artifacts removal is performed by from the combination of detected ICs as artifacts and the clustering of their scalp topography. The clustering of ICs is based on hierarchical clustering with cosine similarity as a distance metric applied on columns of mixing matrix **A**. A threshold decides the separation of subtrees on dendrogram and as a result the number of clusters. If threshold is equal with the largest distance between two existing clusters, the number of clusters is two. The number of clusters is increased as much as the value of threshold becomes smaller than the percentage of biggest distance. The 35% of the biggest distance for separating the dendrogram result to an essential result and the number of subtrees (clusters) being eight. The number of ICs ranges from 3 to 8 ICs on every cluster.
A second application of ICA yields even a better decomposition of brain activity into the available components. This second-stage ICA is applied on the reconstruction of MEG channels from thirty (artifact-free) ICs. Then, detection and elimination of artifacts in the ICs is performed as before, but the values of thresholds of metrics are now different. After removing ten additional artifacts, the number of clusters is six setting 42% of the biggest distance as threshold.

Figure 1 depicts a representative IC for a cluster which is obtained as the mean of the six clustered ICs as in [19]. Following the derivation of representative topographies, a specific location on the head map is computed for each cluster as the centroid of activation of the corresponding representative IC. The ICs used for the estimation of MI and IPDC have common topographies and location as in as in Fig. 1 and 2 but different signal values for each dataset. Overall, six centroids (L) are localized on the scalp topography of Figure 2, which are defined by the maximum value of brain density from every representative ICs (columns of mixing matrix **A** which used for representation of brain density on scalp topography).

We performed a statistical analysis on the estimated values of MI and IPDC. More especially, two-sample t-test and non-parametric Mann-Whitney test are used from the extraction of significant values. Significant diffferences are ploted in Figure 2 using the scalp localization for each cluster showing only the significant diffferences and directions of synchronization.

B. Results

The results of synchronization and statistical analysis are presented in Figure 2 with the specific class topography per cluster.
MI is defined as a non-linear synchronization metric of two time series. Values of MI are calculated separately on every brain rhythm, using a Butterworth band-pass filter of 5th order. Regarding the MI measure, there is significant synchronization difference at L4 and L6 on θ and α_1 bands between the NI and RD groups. In addition, significant differences are observed between L1 and L4, L5 and between L6 and L2, L4 on the α_2 and β bands, respectively.

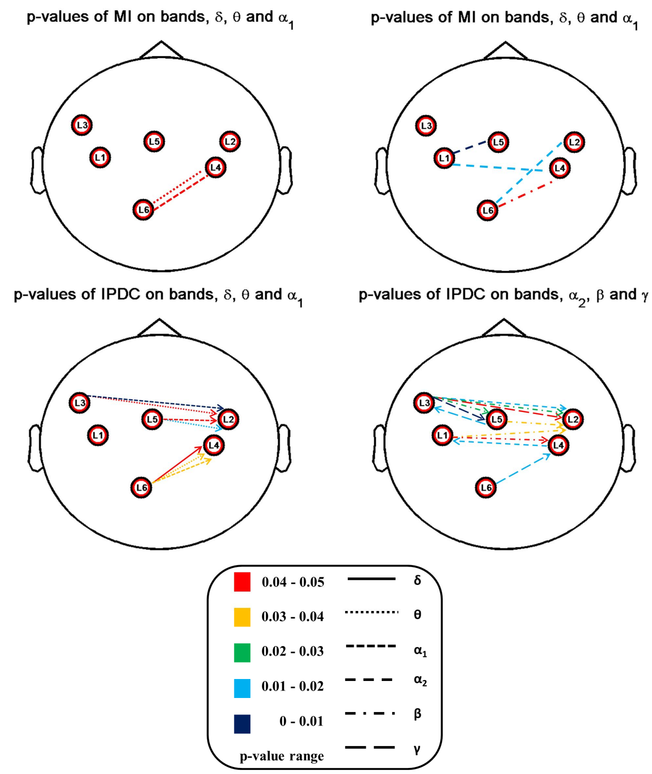

Fig. 2. Significant diffferences for IPDC and MI.

IPDC provides information about the intensity and the directions of information flow of brain networks. Information flow differences are presented in Fig. 2. As it can be observed, a specific connectivity pattern (L3->L2, L6->L4) appears to be implicated in the existence of reading difficulties for both low rhythms (δ, θ, α_1) and high rhythms (α_2, β, γ). Specifically, intense differentiation of information flow is observed from L3 to L2 on θ, α_1, α_2, β, γ brain rhythms, from L6 to L4 on δ, θ, α_1, α_2 brain rhythms, from L5 to L2 on θ, α_1, α_2, β brain rhythms. Furthermore, fewer differences on information flow patterns can be observed for pairs L5->L2 on low frequency bands and for pairs L3->L5, L5->L2, L1->L4, L4->L1, and L1->L2 on high frequency bands.

IV. Discussion

In this paper, coupling synchronization between controls and children with dyslexia from MEG surface signals is investigated. After the identification and rejection of artifact components, an application of clustering ICA algorithm is proposed in order to reduce effectively data dimensionality and to perform the analysis in representative independent data. As it can be proved a difficult task to make conclusions about connectivity patterns among many scalp locations, this procedure can be very useful when one needs to analyze multichannel high resolution MEG signals.

From the analysis of results it can be concluded that IPDC reveals patterns of information flow that is implicated in reading difficulties. These patterns are appeared typically in representative centroids that are located in different head hemispheres. This reflects a communication problem in information processing between the two hemispheres. In addition, children with reading difficulties have reduced IPDC values comparing with controls enforcing the notion of the difficulty transmitting information flows in dyslexia.

Another interesting point of view is that there is a consistent differentiation pattern between low frequency bands (δ, θ, α_1) and high frequency bands (α_2, β, γ). However, in high frequency bands more information flows appear that differentiate the two groups (controls, reading difficulties).

Mutual information shows similar differences and emphasizes the implication of L6-L4 hemisphere-communication in presence of reading difficulties. Also according to this measure, high frequency bands reveal more coupling synchronization patterns implicated in reading difficulties. IPDC, except the additional information of directionality, seems more effective in the prominence of synchronization patterns comparing with mutual information.

Acknowledgment

The authors wish to thank Andrew C. Papanicolaou, Department of Pediatrics, University of Texas and Panagiotis G. Simos, Department of Psychology, University of Crete about the dataset of control and dyslexic group.

References

[1] D.S. Bassett, E.T. Bullmore, Human brain networks in health and disease, Curr Opin Neurol, vol. 22 (4), pp. 340-347, Aug 2009.

[2] C. Jin, C. Gao, C. Chen, S. Ma, R. Netra, Y. Wang, M. Zhang. D. Li, A preliminary study of the dysregulation of the resting networks in first-episode medication-naïve adolescent depression, Neurosci Lett, vol. 503(2), pp. 105-109, 3 Oct 2011.

[3] C. J. Stam, , W. de Haan, A. Daffertshofer, B. F. Jones, I. Manshanden , A. M. vanCappellen van Walsum, T. Montez, J. P. Verbunt, J. C. de Munck, B. W. vanDijk, H. W. Berendse, P. Scheltens, Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease, Brain, vol. 132(Pt 1), pp. 213-14,Jan 2009.

[4] D. S. Basset, , B. G. Nelson, B. A. Mueller, , J. Camchong, and K. O. Lim, Altered resting state complexity in schizophrenia, Neuroimage, vol. 59(3), pp. 2196-207, 1 Feb 2012 .

[5] S. I. Dimitriadis, N. A. Laskaris, P. G. Simos, S. Micheloyannis, J. M. Fletcher, R. Rezaie, and A. C. Papanicolaou, Altered temporal correlations in resting-state connectivity fluctuations in children with reading difficulties detected via MEG, Neuroimage, vol. 83, pp. 307-317, Dec 2013.

[6] V. A. Grin-Yatsenko, I. Baas, V. A. Ponomarev, J. D. Kropotov, Independent component approach to the analysis of EEG recordings at early stages of depressive disorders, Clin Neurophysiol, vol. 121(3), pp. 281-9, Mar 2010.

[7] V. Tsiaras, P. G. Simos, R. Rezaie, B. R. Sheth, E. Garyfallidis, E. M. Castillo, A. C. Papanicolaou, , Extracting biomarkers of autism from MEG resting-state functional connectivity networks, Comput Bio Med, vol. 41(12), pp. 1166-77, Dec 2011.

[8] R. Quian Quiroga, A. Kraskov, T. Kreuz, P. Grassberger, Performance of different synchronization measures in real data: a case study on electroencephalographic signals, Phys Rev E. Stat Nonlin Soft Matter Phys, vol. 65(4 Pt 1), pp. 041903, Apr 2002.

[9] L.A. Baccala, K. Sameshima, Partial directed coherence: a new concept in neural structure determination, Biol Cybern, vol. 84(6), pp. 463-74, Jun 2001.

[10] R. Kuś, M. Kamiński, K.J. Blinowska, Determination of EEG activity propagation: pair-wise versus multichannel estimate, IEEE Trans Biomed Eng, vol. 51(9), pp. 1501-10, Sep 2004.

[11] R. Rezaie, P. G. Simos, J. M. Fletcher, J. Juranek, P. T. Cirino, Z. Li, A. D. Passaro and A. C. Papanicolaou, The timing and strength of regional brain activation associated with word recognition in children with reading difficulties, Front Hum Neurosci, pp. 5:45, 2011.

[12] A. J. Bell, T. J. Sejnowski, An information-maximization approach to blind separation and blind deconvolution, Neural Comput, vol. 7(6), pp. 1129-59, Nov 1995.

[13] L. E. Ethridge, S. M. Malone, W. G. Iacono, B. A. Clementz, Genetic influences on composite neural activations supporting visual target identification, Biol Psychol, vol. 92(2), pp.329-41, Feb 2013.

[14] R. Oostenveld, P. Fries, E. Maris and J. M. Schoffelen, FielTrip: Open source software for advantage analysis of MEG, EEG, and invasive electrophysiological data, Comput Intell Neurosci, vol. 2011, pp. 156869, 2011.

[15] J. Escudero, R. Hornero, D. Abásolo, A. Fernández, Quantitative evaluation of artifact removal in real magnetoencephalogram signals with blind source separation, Ann Biomed Eng, vol. 39(8), pp. 2274-86, Aug 2011.

[16] G. Barbati, C. Porcaro, F. Zappasodi, P. Maria Rossini, F. Tecchio, Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals, Clin Neurophysiol, vol.115(5), pp.1220-32, May 2004.

[17] J. P. Lindsen, J. Bhattacharya, Correction of blink artifacts using independent component analysis and empirical mode decomposition, Phychophysiology, vol. 47(5), pp. 955-60, Sep 2010.

[18] R.T. Rato, M.D. Ortigueira, A.G. Batista, On the HHT, its problems, and some solutions, Mech Sys and Sig Proc , vol. 22(6), pp. 1374-94, Aug 2008.

[19] R. Grandchamp, C. Braboszcz, S. Makeig, A. Delorme, Stability of ICA decomposition across within-subject EEG datasets, Conf Proc IEEE Eng Med Biol Soc, vol. 2012, pp. 6735-9, 2012.

[20] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning*, 2nd ed. vol. 14.3.12, pp.520-528, 2009.

[21] "The DISTANCE Procedure: Proximity Measures". SAS/STAT 9.2 Users Guide. SAS Institute. Retrieved 2009-04-26

[22] J. Onton, M. Westerfield, J. Townsend, S. Makeig, Imaging human EEG dynamics using independent component analysis, Neurosci Biobehav Rev, vol. 30(6), pp. 808-22, 2006.

[23] W. Klimesch, EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis, Brain Res Brain Res Rev, vol. 29(2-3), pp. 169-95, Apr 1999.

[24] S.L. Marple, "Digital Spectral Analysis with Applications", Prentice Hall, 1987.