6th Annual International IEEE EMBS Conference on Neural Engineering
San Diego, California, 6 - 8 November, 2013

Investigating the effective brain networks related to working memory using a modified directed transfer function

Hossein Shahabi, *Student Member*, *IEEE*, Sahar Moghimi, *Member*, *IEEE*, and Ali Moghimi

Abstract— **We utilized a combination of orthogonal least squares (OLS) algorithm and short-time direct directed transfer function (SdDTF) for evaluating the effective brain networks. Directional changes in connectivity among brain regions were assessed during two states, resting and working memory (WM). Our results demonstrated an increase in the strength of connectivity between frontal and parietal regions and a decrease in the strength of connectivity between the regions of the default mode network, during task engagement. Connectivity was prominent in the low frequency bands, theta and alpha. Furthermore, during the WM task, theta oscillations increased in the frontal lobe while a reduction was observed in the alpha band.**

I. INTRODUCTION

The term working memory (WM) refers to a part of the memory system, which is responsible for temporary storage and manipulation of information during complex cognitive tasks (e.g. certain types of learning and reasoning) [1]. Many researchers have employed different approaches to investigate the role of brain regions in a WM task [2]-[8]. For instance, EEG studies pursued the task related dynamic changes in different frequency bands of signals from subject's scalp [2]-[4]. Furthermore, the effect of task difficulty, which is believed to influence the activation of brain resources has been investigated in the context of EEG studies [2], [3]. It is widely reported that theta oscillations, specifically in the frontal lobe, increase with more cognitive effort [2], [7], [8]. In contrast, to the best of our knowledge, there is no common agreement on the pattern of alpha oscillations; some studies reported an increase in the spectral power [4], while others demonstrated a reduction in this band [7]. Investigations on the higher frequency bands have revealed an increase [5], or decrease [8] in the gamma band activity during WM tasks. Exploring the causal interaction of brain regions during a WM task may lead to a better comprehension of the brain networks of the WM system. In recent years, the investigation of brain networks and connectivity between different regions has become an attractive topic for many researchers. Among the connectivity measures applied to brain signals, directed transfer function (DTF) [9] has been repeatedly adopted [10], [11]. DTF is based on multivariate autoregressive (MVAR) modeling of signals in different channels. By performing tests on simulated and real data, Blinowska et al. have argued that the multivariate approach is more reliable than bivariate modeling, since it is capable of identifying the true causal interactions [12]. To overcome the major drawback of DTF, which is its inefficiency to distinguish between direct and indirect connectivity, Korzeniewska introduced the direct DTF (dDTF) [13], which was later extended to short-time direct DTF (SdDTF) for evaluation of the dynamics of nonstationary signals [14]. This method has proven to be capable of indicating the time varying connectivity maps. The current study aims to use EEG to elucidate the directed connectivity among brain regions that are involved in WM. For this purpose, we employed the SdDTF method. The identification algorithm and processing of trial information in cognitive tasks were modified in order to improve the obtained results.

H. Shahabi (e-mail: h.shahabi@ieee.org) and S. Moghimi (corresponding author, e-mail: s.moghimi@um.ac.ir) are with the Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran.

A. Moghimi (e-mail: moghimi@um.ac.ir) is with the Department of biology, Ferdowsi University of Mashhad, Mashhad, Iran.

II. MATERIALS AND METHODS

A. Subjects

9 healthy university students (all male and right-handed; mean age 22.5±3.2) with normal or corrected-to-normal vision and no history of psychiatric or neurological disorder participated in this study. Data from 5 subjects were discarded due to the low quality of signals in some channels. All subjects signed the consent form and were informed about the task procedure.

B. Stimuli

We used the N-back task for evaluating the cognitive effort. One of this task's attractive properties is that by increasing the memory load, the speed and type of visual input remain unchanged [3]. 3 conditions of this task were used, namely 0-back, 1-back and 2-back. In each condition a sequence of letters appeared consecutively. We used 9 English letters (excluding vowels) for designing the experiment. Each letter was presented for 500ms and then a 2000ms inter-stimulus interval with a fixation cross appeared on the screen [3]. During the N-back task participants were instructed to press a key when the current letter matched the Nth previously presented letter. In the 0-Back condition participants pressed a key when 'X' appeared. Subjects performed the tasks in 3 sessions. Each session consisted of 4 blocks: 3 N-back and 1 baseline block, which were presented randomly. In each N-back block 30 letters (10 targets) were presented and therefore it lasted 75secs. Similarly we set the length of baseline block to 75secs.

C. EEG recording

EEG data were recorded with 31 electrodes (Mitsar-EEG-202 system): international 10/20 electrodes plus 10

978-1-4673-1969-0/13/$31.00 ©2013 IEEE

interpolated electrodes. The EEG signals were sampled at 250 Hz and referenced to the average of left and right ears.

D. Short time direct directed transfer function

The directed transfer function is an explicit function of multivariate autoregressive (MVAR) model parameters [9]. By assuming $\mathbf{x}(t) = [x_1(t)\ x_2(t)\ \ldots\ x_L(t)]^T$ to include the measurements of L channels in time t, the MVAR process can be expressed as [14]:

$$\sum_{r=0}^{p} \mathbf{A}_r \mathbf{x}(t-r) = \mathbf{e}(t) \qquad (1)$$

where p is the model order and $\mathbf{e}(t)$ is a zero mean uncorrelated residual noise vector. $\mathbf{A}_0$ is an $L\times L$ negative identity matrix, and $\mathbf{A}_r$ is an $L\times L$ MVAR coefficients matrix which contains the weights of input delays for output construction [14]. Computing the Fourier transform of (1) will result in

$$\mathbf{A}(f)\mathbf{X}(f) = \mathbf{E}(f) \qquad (2)$$

with

$$\mathbf{A}(f) = \sum_{r=0}^{p} \mathbf{A}_r e^{-j2\pi r f \Delta t}$$

f is frequency and Δt is the sampling interval. From (2) the transfer function of the multi-input system, $\mathbf{H}(f)$ which is the inverse of matrix $\mathbf{A}(f)$ can be defined

$$\mathbf{X}(f) = \mathbf{H}(f)\mathbf{E}(f) \qquad (3)$$

h_{mn} is an element of the transfer matrix $\mathbf{H}(f)$ and defines the transfer function between the m-th output and the n-th input of the system. Therefore, if the absolute value of h_{mn} is significantly different from zero, the information in channel n can be used for estimating channel m. By assuming $\mathbf{V}$, as the estimated residual noise covariance matrix, we can obtain the spectral matrix [14]

$$\mathbf{S}(f) = \mathbf{H}(f)\mathbf{V}\hat{\mathbf{H}}(f) \qquad (4)$$

where '^' shows the conjugate-transpose of matrix $\mathbf{H}(f)$. Next we suppose $\mathbf{C}(f)$ as the inverse of spectral matrix $\mathbf{S}(f)$, with its values called c_{mn}. As a result, the partial coherence may be defined in the form,

$$\delta_{mn}(f) = \frac{c_{mn}(f)}{\sqrt{c_{mm}(f)c_{nn}(f)}} \qquad (5)$$

The Combination of partial coherence and the transfer matrix $\mathbf{H}(f)$ is the SdDTF [14],

$$\lambda_{mn} = \frac{|h_{mn}(f)||\delta_{mn}(f)|}{\sqrt{\sum_f \sum_{mn} |h_{mn}(f)|^2 |\delta_{mn}(f)|^2}} \qquad (6)$$

This new index eliminates the problem of indirect connectivity, which was neglected in DTF. Furthermore, it has been illustrated that this new approach for normalization gives more accurate results in comparison with DTF or dDTF [14]. The complete description of SdDTF is available in [14]. We used (6) for producing the connectivity matrices.

E. Data analysis

After recording the EEG signals, the 3 sessions were divided into 4 groups, namely baseline, 0-back, 1-back, and 2-back. Each group contained 90 trials. All signals were assessed completely and trials with artifact were rejected. Next a high-pass filter, with its cutoff frequency located at 0.5Hz was applied to all non-artifact signals. Although all the tasks were performed thoroughly, here we only report and compare the results obtained by analyzing the baseline and 2-back conditions. The non-target trials were used for the analysis of the 2-back condition, in order to remove the motor activity from the WM networks. Moreover, the number of non-target trials were two times larger than the number of target trials. By considering the subjects' performance, the effect of motor activity in non-target trials was assumed negligible.

A common choice for calculating the MVAR model coefficients is the Yule-Walker algorithm. However, we employed the orthogonal least squares (OLS) algorithm [15], where only a limited number of the candidate regressors, different delays of all channels (p × L regressors), were chosen according to their similarity to the modeled signal. By applying this procedure the majority of the elements of $\mathbf{A}_r$ in (1) were set to zero, resulting in a sparse MVAR model. The Akaike information criteria (AIC) was used for selecting the value of p in (1). We selected p=8 for all subjects, since the slope of the AIC/p diagram was close to zero for this value. For the sake of comparison between the different developed models, a similar value of p was selected for all subjects. We applied the OLS method to each trial and then estimated the connectivity matrix for each trial separately, therefore considering the dynamic and nonstationary nature of the system. The window length for calculating the SdDTF index was 2500ms. The trial related connectivity matrices were averaged in the final step to build a general matrix. Using this matrix we investigated the connectivity patterns in different frequency bands.

III. Results

In order to highlight the differences in the connectivity patterns of baseline and 2-back conditions, we subtracted the 2-back related connectivity matrix from that of the baseline for each subject. Since an exactly similar connectivity pattern between the involved brain regions cannot be expected among all the participants, here we report the obtained results for each subject separately. Two frequency bands, theta (4-7 Hz) and alpha (8-12 Hz), mostly discussed in WM literature, were selected. Among the $L\times L - L$ connectivity coefficients only those at or above the 98th percentile are reported. The activity and connectivity maps are illustrated in Fig. 1. Differences in the connectivity patterns of 2-back and baseline conditions in theta and alpha bands are presented in Fig. 1A and Fig. 1C for different subjects. As a result of task engagement the interconnectivity in the prefrontal cortex, as well as the long range connectivity from frontal to parietal and parietal to frontal lobes increased in most subjects. These findings are

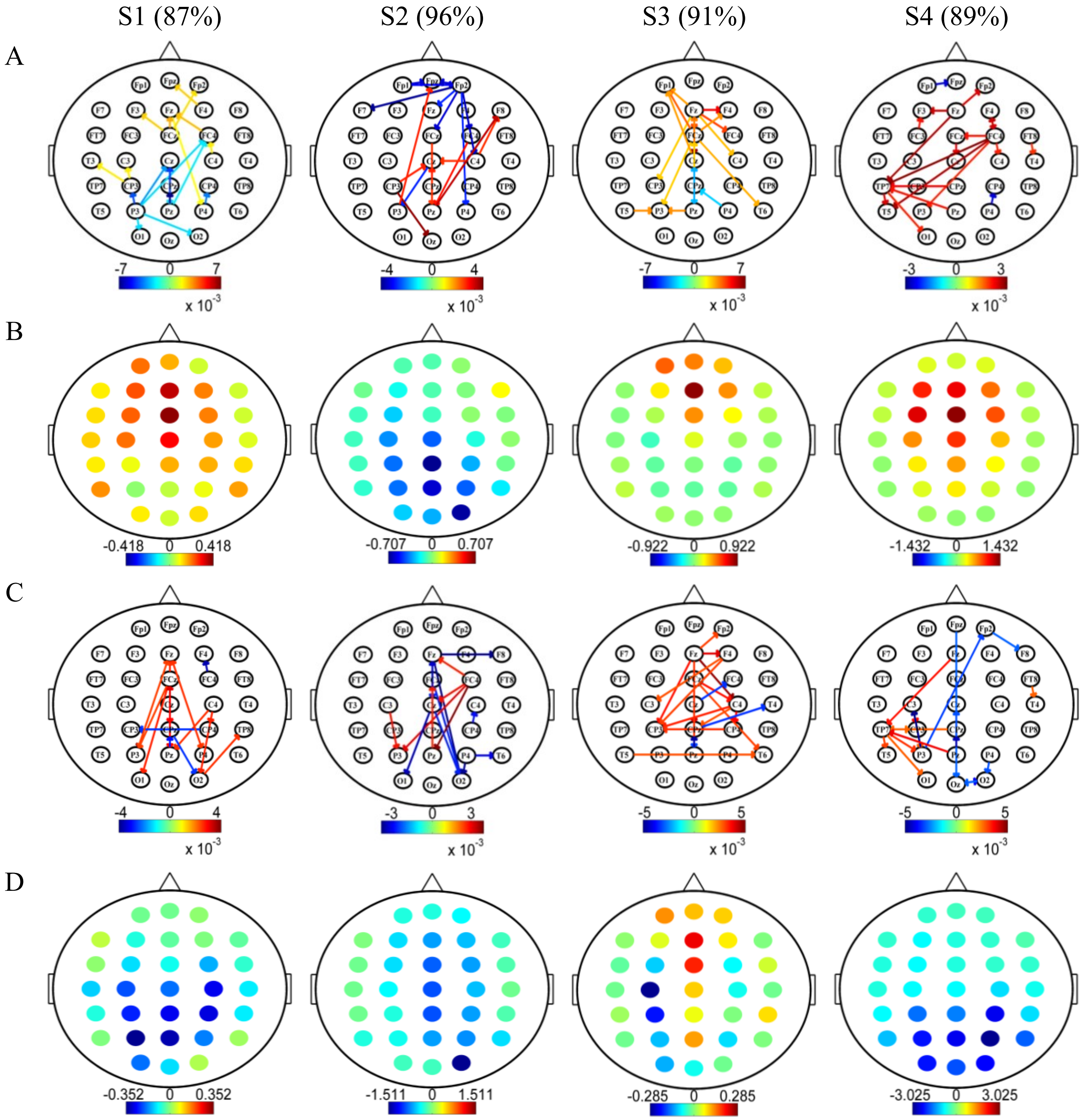

Figure 1. Difference in brain activty and connectivity for 4 subjects between 2-back and baseline conditions. (A) Connectivity maps in the theta band. (B) Activity of theta band. (C) Connectivity map in the alpha band. (D) Activity of alpha band. Numbers in parantheses are the overall performance of each subject in the 2-back task.

consistent with a previous EEG study of WM that indicated the role of frontal and parietal lobes and their associations in WM tasks [16]. Recently an fMRI study of the WM network during an N-back task identified the following 5 most prominent anatomical substrates of the WM system: right middle frontal gyrus and frontal pole, Right and left lateral occipital cortex and precuneus/parietal lobe [17]. Our results show a significant increase in connectivity between these regions. Although we observed increased connectivity between different regions of the WM network, reduced connectivity was also experimented between some of the nodes (Fig. 1A & C). This reduction in connectivity between some brain areas can be explained by considering the deactivation of the regions involved in the default mode network (DMN) during WM tasks [18]. The DMN regions mainly comprise precuneus, medial frontal, inferior parietal cortical regions and medial temporal lobe, and it is hypothesized that these regions have more activity and connectivity in the resting state than during cognitive tasks [18]. Our results also show that the interconnectivity and long range connectivity between regions of the DMN are suppressed during the 2-back task. Theta band activity which is widely discussed in the WM literature is illustrated in Fig. 1B. In most participants frontal theta increased in the 2-back task in comparison to the baseline condition. This is in agreement with the previous studies [2], [7], [8]. Analysis of the alpha band revealed a significant reduction, specifically in the parietal regions, which was also reported by other researchers [7].

IV. Conclusion and Discussion

In this paper we investigated the directed causal interactions between brain sites during WM tasks by utilizing a combination of the OLS and SdDTF algorithms. The required codes for implementing the methods were developed in MATLAB without using any toolbox. Moreover, we extracted the maps of brain oscillations in different frequency bands. Analyzing the non-target trials (instead of target trials), resulted in removal of motor effects from calculation of the MVAR model coefficients. The results were consistent with the previous findings regarding the WM studies.

As explained in the previous section, DMN and WMN had overlaps in some regions and both illustrated changes during a WM task. It is plausible to assume that simultaneous assessment of these networks can provide valuable information about the neural correlates of the phenomenon under study.

Calculation of the MVAR model coefficients is the key aspect of directed connectivity studies. In this study, we used the OLS method, to obtain an accurate and sparse model. This algorithm may be more suitable for investigating the small-world properties of brain networks. The small-world properties have been defined for graphs with specific features [19]. One of these properties states that there are only few direct connections between nodes of a graph. It is suggested that structural and functional brain networks have the small-world properties, meaning that each region of the brain is not directly connected to all other regions or influenced by them [19]. The OLS identification method provides a sparse connectivity pattern, since it utilizes the dominant regressors (time-series) for estimating each channel. We observed that the selected channels were mainly the neighbor channels with the exception of only a few distant channels. This result, which was due to the application of the OLS method was in agreement with the mentioned property of small-world networks.

Acknowledgment

The authors would like to thank the staff at AREN institute for their assistance during the EEG recording sessions.

References

[1] A. Baddeley, “Working memory: Looking back and looking forward,” *Nature Reviews Neuroscience*, vol. 4, no. 10, pp. 829-839, 2003.

[2] L. Michels, K. Bucher, R. Lüchinger, P. Klaver, E. Martin, D. Jeanmonod, and D. Brandeis, “Simultaneous EEG-fMRI during a working memory task: modulations in low and high frequency bands,” *PLoS One*, vol. 5, no. 4, e10298, 2010.

[3] A. Brouwer, M. A. Hogervorst, J. BF V. Erp, T. Heffelaar, P. H. Zimmerman, and R, Oostenveld, “Estimating workload using EEG spectral power and ERPs in the n-back task,” *Journal of Neural Engineering*, vol. 9, no. 4, 045008, 2012.

[4] O. Jensen, J. Gelfand, J. Kounios, and J. E. Lisman, “Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task,” *Cerebral Cortex*, vol. 12, no. 8, pp. 877-882, 2002.

[5] M. W. Howard, D. S. Rizzuto, J. B. Caplan, J. R. Madsen, J. Lisman, *et al*., “Gamma oscillations correlate with working memory load in humans,” *Cerebral Cortex*, vol. 13, no. 12, pp. 1369-1374, 2003.

[6] C. E. Curtis and M. D'Esposito, “Persistent activity in the prefrontal cortex during working memory,” *Trends in cognitive sciences*, vol. 7, no. 9, pp. 415-423, 2003.

[7] A. Gevins, M. E. Smith, L. McEvoy, and D. Yu, “High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice,” *Cerebral Cortex*, vol. 7, no. 4, pp. 374-385, 1997.

[8] M. J. Brookes, J. R. Wood, C. M. Stevenson, J. M. Zumer, T. P. White, *et al*., “Changes in brain network activity during working memory tasks: a magnetoencephalography study,” *Neuroimage*, vol. 55, no. 4, pp. 1804-1815, 2011.

[9] M. Kamiński, M. Ding, W. A. Truccolo, and S. L. Bressler, “Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance,” *Biological cybernetics*, vol. 85, no. 2, pp. 145-157, 2001.

[10] F. Babiloni, F. Cincotti, C. Babiloni, F. Carducci, D. Mattia, *et al*., “Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function,” *Neuroimage*, vol. 24, no. 1, pp. 118-131, 2005.

[11] C. Babiloni, F. Vecchio, S. Cappa, P. Pasqualetti, S. Rossi, *et al*., “Functional frontoparietal connectivity during encoding and retrieval processes follows HERA model: A high-resolution study,” *Brain Research Bulletin*, vol. 68, no. 4, pp. 203-212, 2006.

[12] K. J. Blinowska, R. Kuś, and M. Kamiński, “Granger causality and information flow in multivariate processes,” *Physical Review E*, vol. 70, no. 5, 050902, 2004.

[13] A. Korzeniewska, M. Mańczak, M. Kamiński, K. J. Blinowska, and S. Kasicki, “Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method,” *Journal of neuroscience methods*, vol. 125, no. 1, pp. 195-207, 2003.

[14] A. Korzeniewska, C. M. Crainiceanu, R. Kuś, P. J. Franaszczuk, and N. E. Crone, “Dynamics of event-related causality in brain electrical activity,” *Human brain mapping*, vol. 29, no. 10, pp. 1170-1192, 2008.

[15] S. A. Billings, S. Chen, and M. J. Korenberg, “Identification of MIMO non-linear systems using a forward-regression orthogonal estimator,” *International Journal of Control*, vol. 49, no. 6, pp. 2157-2189,1989

[16] P. Sauseng, W. Klimesch, M. Schabus, and M. Doppelmayr, “Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory,” *International Journal of Psychophysiology*, vol. 57, no. 2, pp. 97-103, 2005.

[17] R. Sala-Llonch, C. Peña-Gómez, E. M. Arenaza-Urquijo, D. Vidal-Piñeiro, N. Bargalló, *et al*., “Brain connectivity during resting state and subsequent working memory task predicts behavioural performance,” *Cortex*, vol. 48, no. 9, pp. 1187-1196, 2012.

[18] M. P. Van Den Heuvel, and H. E. Hulshoff Pol, “Exploring the brain network: a review on resting-state fMRI functional connectivity,” *European Neuropsychopharmacology*, vol. 20, no. 8, pp. 519-534, 2010.

[19] E. Bullmore and O. Sporns, “Complex brain networks: graph theoretical analysis of structural and functional systems,” *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186-198, 2009.