35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

A Tensorial Approach to Access Cognitive Workload related to Mental Arithmetic from EEG Functional Connectivity Estimates

S.I. Dimitriadis, Yu Sun, K. Kwok, N.A. Laskaris, A. Bezerianos*

Abstract— **The association of functional connectivity patterns with particular cognitive tasks has long been a topic of interest in neuroscience, e.g., studies of functional connectivity have demonstrated its potential use for decoding various brain states. However, the high-dimensionality of the pairwise functional connectivity limits its usefulness in some real-time applications. In the present study, the methodology of tensor subspace analysis (TSA) is used to reduce the initial high-dimensionality of the pairwise coupling in the original functional connectivity network to a space of condensed descriptive power, which would significantly decrease the computational cost and facilitate the differentiation of brain states. We assess the feasibility of the proposed method on EEG recordings when the subject was performing mental arithmetic task which differ only in the difficulty level (easy: 1-digit addition *v.s.* 3-digit additions). Two different cortical connective networks were detected, and by comparing the functional connectivity networks in different work states, it was found that the task-difficulty is best reflected in the connectivity structure of sub-graphs extending over parieto-occipital sites. Incorporating this data-driven information within original TSA methodology, we succeeded in predicting the difficulty level from connectivity patterns in an efficient way that can be implemented so as to work in real-time.**

I. INTRODUCTION

Recent advances in neuroengineering have made possible the use of EEG signals or "brain waves" for communication between humans and computers. There is a rapidly growing interest in Brain Computer Interface (BCI) and human-machine interaction (HMI) areas for methodologies that will not only provide communication and control capabilities to people with severe motor disabilities, but also will facilitate novel applications for healthy subjects (e.g. cognitive monitoring [1]). This is a multi-disciplinary research area in which engineers and computer scientists interact with experts from medicine, neuroscience and psychology. Any practical implementation of a BCI design requires an efficient signal processing scheme that includes signal conditioning, feature extraction and classification. In this paper, we focus on the feature extraction step and introduce the use of a tensorial algorithm that learns, in an intelligent way, the essence of original, high-dimensional domain of pairwise couplings. The proposed algorithm reduces dramatically the necessary efforts in the subsequent classification step, which now is performed efficiently in a space of lower dimensionality and with minimal computational complexity.

Most of the previous research works in brain decoding represented functional connectivity graphs (FCGs) as *vectors* in a high-dimensional space (e.g. [2-4]). The main drawback of such an approach (and the related feature extraction algorithms) is that it overlooks the inherent format of FCGs. Each FCG has a natural tabular representation and hence can be thought of as second-order tensor. The relationship between the row and column vectors of the associated matrix might be important for deriving a suitable low-dimensional representation (projection), especially when the number of available connectivity patterns is small. To overcome this limitation, we treat FCGs as tensors and resort to tensor subspace analysis (TSA) [5] as the most appropriate feature extraction algorithm.

The principal scope of this work was to exploit the recently introduced subspace learning algorithm of TSA and suggest algorithmic strategies that will treat the functional connectivity estimates directly (i.e. without computing network metrics), and, mainly, at the level of individual trials. In our approach, we treated the trial-dependent functional connectivity matrix as a tensor and used the TSA analysis to project it in a lower-dimensionality space where assessments of different cognitive loads would be more efficient. The methodological advantage relies on the concept that once the mapping (from trial-based functional connectivity domain to a reduced space) is learned, its application to data from an unseen trial will only require (apart from the connectivity estimates) a few matrix multiplications. To provide "a proof of concept"', we describe below the algorithmic steps taken (and the related evaluation) for discriminating between an easy and difficult arithmetic task.

II. SUBSPACE LEARNING BASED ON TENSOR ANALYSIS

In this section, we present a new algorithm for FCG representation based on the considerations of multi-linear

S.I. Dimitriadis and N.A. Laskaris is with the Artificial Intelligence and Information Analysis Laboratory, Department of Informatics, Aristotle University, Thessaloniki, 54124, Greece, (e-mail: stidimitriadis@gmail.com; laskaris@aiia.csd.auth.gr).

Y. Sun and A. Bezerianos is with the SINAPSE, National University of Singapore, 28 Medical Drive, 117456, Singapore, (e-mail: lsisu@nus.edu.sg; tassos.bezerianos@nus.edu.sg). *Asterisk indicates corresponding authors.*

K. Kwok is with Temasek Laboratories, National University of Singapore, 117411, Singapore, (e-mail: kenkwok@nus.edu.sg).

algebra and differential geometry. Given some FCGs sampled from the functional connectivity manifold, we can build an adjacency graph to model the local geometrical structure of the data manifold. TSA derives a projection that respects this graph structure. The obtained tensor subspace provides an optimal linear approximation to the FCG manifold.

A. The linear Dimensionality Reduction Problem in Tensor Space

Let $X \in \Re^{m_1 \times m_2}$ be a FCG of size $m_1 \times m_2$. Mathematically, X can be thought as a 2nd order tensor (or 2-tensor) in the tensor space $\Re^{m_1} \otimes \Re^{m_2}$. The generic problem of linear dimensionality reduction in the second order space is the following. Given a set of tensors (i.e. matrices) $X_1,..., X_m \in \Re^{n1} \otimes \Re^{n2}$ find two transformation matrices U of size $n_1 \times l_1$ and V of $n_2 \times l_2$ that maps these m tensors to a set of tensors $Y_1,...,Y_m \in \Re^{l1} \otimes \Re^{l2} (l_1 < n_1, l_2 < n_2)$, such that Y_i "represents" X_i, where $Y_i = U^T X_i V$. The method is of particular interest in the special case where $X_1,..., X_m \in M$ and M is a nonlinear sub-manifold embedded in $\Re^{n_1} \otimes \Re^{n_2}$.

B. Optimal Linear Embedding

The domain of FCGs is most probably a nonlinear sub-manifold embedded in the tensor space. With the adopted TSA, we sought to estimate geometrical and topological properties of the sub-manifold from "scattered data" (i.e. randomly selected connectivity matrices) lying on this unknown sub-manifold. We will consider the particular question of finding a linear subspace approximation to the sub-manifold in the sense of local isometry. TSA is fundamentally based on Locality Preserving Projection (LPP) [6].

Given m data points $X = \{X_1,..., X_m\}$ sampled from the FCG sub-manifold $M \in \Re^{n_1} \otimes \Re^{n_2}$, one can build a nearest neighbor graph G to model the local geometrical structure of M. Let S be the weight matrix of G. A possible definition of S is as follows:

$$S_{ij} = \begin{cases} e^{-\|X_i - X_j\|^2 / t} & \text{If } X_j \text{ is among the k nearest neighbors of } X_i \\ 0 & \text{Otherwise} \end{cases} \tag{1}$$

where t is a suitable constant. The function $e^{-\|X_i - X_j\|^2 / t}$ is the so called heat kernel which is intimately related to the manifold structure. $\|\cdot\|$ is the Frobenius norm of matrix. In the case of supervised learning (classification labels are available), the label information can be easily incorporated into the graph as follows:

$$S_{ij} = \begin{cases} e^{-\|X_i - X_j\|^2 / t} & \text{If } X_j \text{ and } X_i \text{ share the same label;} \\ 0 & \text{Otherwise} \end{cases} \tag{2}$$

Let U and V be the transformation matrices. A reasonable transformation respecting the graph structure can be obtained by solving the following objective functions:

$$\min_{U,V} \sum_{ij} \|U^T X_i V - U^T X_j V\|^2 S_{ij} \tag{3}$$

The objective function incurs a heavy penalty if neighboring points X_i and X_j are mapped far apart. With D be a diagonal matrix $D_{ii} = \sum_j S_{ij}$ and after mathematical calculations (for further details see [3]), the optimization problem was restricted to the following equations:

$$(D_U - S_U)V = \lambda D_U V \tag{4}$$

Once V is obtained, U can be updated by solving the following generalized eigenvector problem:

$$(D_V - S_V)V = \lambda D_V V \tag{5}$$

Thus, the optimal U and V can be obtained by iteratively computing the generalized eigenvectors of (4) and (5).

III. EXPERIMENTAL DATA (FROM EEG RECORDINGS TO TRIMMED FCGS)

A. Experimental task

We analyzed functional connectivity patterns derived from EEG recordings of one subject (male) performing a mental arithmetic task (addition) with six levels of difficulty. EEG data was recorded from 64 channels at 256 Hz with an ActiveTwo Biosemi systm and referenced using average reference. The experiment was segmented into blocks of one minute, with a thirty second rest period between each block. Within a block (labeled as different levels, e.g., Lvl3 & Lvl5 in Fig. 1), all problems were of the same difficulty level (0 to 5). A schematic diagram of the experimental protocol is presented here:

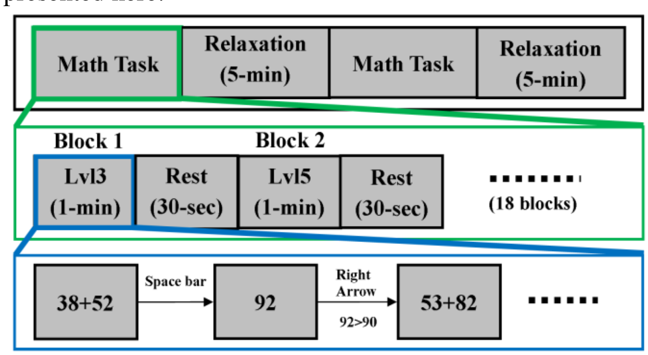

Fig.1. Sequence of events in the mental arithmetic task.

The details of the experiments have been described previously [7]. Briefly, arithmetic summation problems consisting of two figures were displayed continuously until the subject pressed the space bar to indicate that he had obtained the answer. The problem was then replaced by a proposed answer that was either equal to, larger than, or smaller than the true answer, with a probability of 33% for each case. The participant then indicated which one of the three possibilities was the correct one by pressing the corresponding arrow key (down for equal, right for larger, and left for smaller). Difficulty levels were randomized with the constraint that all difficulty levels were presented an equal number of times and that adjacent blocks were of a different difficulty level. After three repetitions of the six difficulty levels, the participant had a 5-minute relaxation break comprising a slide show of landscape pictures.

B. Preprocessing

Using the trigger-information about when the numbers were presented on the screen and the user's reaction time stamp, we segmented the signals into trials (of variable length; starting 100 ms after the onset of the stimulus and lasting up to 200 ms before button press). Biological artifacts were diminished by means of ICA [8, 9] employing function *runica* from EEGLAB [10]. Signals were filtered within frequency range of 0.5 to 45 Hz (from δ to γ band). This study intended to introduce a TSA based FCG analysis method and investigate its sensitivity in differentiating between different cognitive workload. Therefore, the data from level 1 (addition of 1-digit numbers) and level 5 (addition of 3-digit numbers) was used for the following investigation.

C. Initial FCG Derivation and standard network analysis

Since we have recently showed that phase synchrony contributes significantly to the neural substrate of mental calculations [11], we commenced by constructing functional connectivity graphs using a phase coupling estimator called PLV [8, 11, 12], for each trial, condition and frequency bands. We then computed the Local Efficiency (denoted hereafter as LE [8]). The LE was defined as:

$$LE = \frac{1}{N}\sum_{i\in N}\frac{\sum_{j,h\in G_{i,j,h\neq i}}(d_{jh})^{-1}}{k_i(k_i-1)} \qquad (6)$$

wherein k_i corresponded to the total number of spatial directed neighbors (first level neighbors) of the current node, N was the set of all nodes in the network, and d set the shortest absolute path length between every possible pair in the neighborhood of the current node.

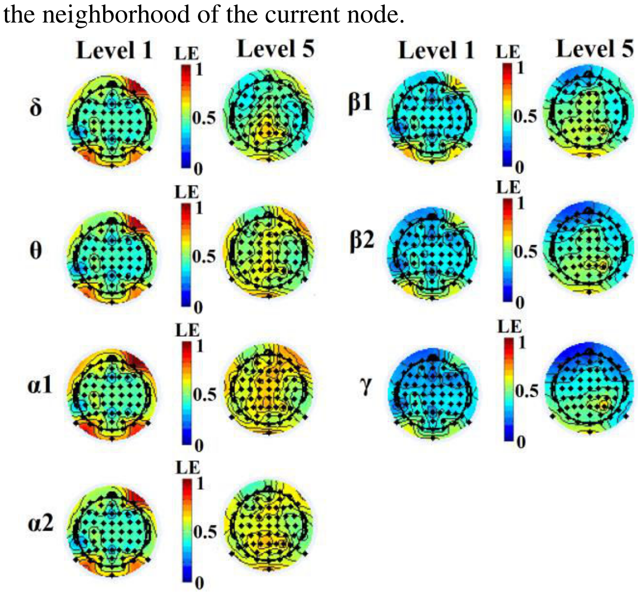

Fig.2. Brain topographies of nodal^{LE} between tasks of level 1 and level 5.

LE is understood as a measure of the fault tolerance of the network, indicative of how well sub-graphs exchange information when the indexed node is eliminated [13, 14]. Specifically, each node was assigned the shortest path length within its sub-graph, G_i. In order to reduce the computational load of TSA, we exploited the differences seen in nodal^{LE} measurements (expressing information exchange rate in the vicinity of each node), between the two extreme conditions (level 1 *v.s.* level 5). Fig.2 compares topographically these measurements across frequency bands and (in accordance with related literature) revealed two sensors (PO7 and PO8) that differ systematically – across frequency bands – between the two difficulty levels.

Fig. 3 shows the averaged values of LE^{PO7} & LE^{PO8} across frequency bands in both levels. Statistical significance of the differences in LE^{PO7} & LE^{PO8} associated with Level 1 and Level 5 was assessed in each frequency band independently ($p < 0.0001/7$, Bonferroni corrected).

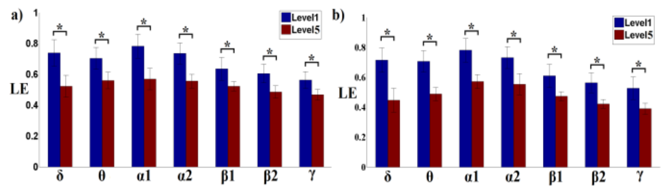

Fig.3. Trial–averaged nodal^{LE} measurements (across various frequency bands) for difficulty Level 1 and 5 for a) PO7 and b) PO8 sensors ($p < 0.0001/7$, Wilcoxon Test, Bonferroni corrected).

D. Trimmed FCGs

We reduced the input for tensor analysis from FCGs of dimension $N \times N$ (where $N = 64$ denotes the original number of sensors) to sub-graphs of size $N' \times N'$ (with $N' = 12$ sensors corresponding to the spatial neighborhood of PO7 and PO8 sensors). Each connection can be seen as feature used in TSA. The total number of connections between sensors belonging in the neighborhood is: 66/2016≈3% of the total number of possible connections in the original graph. The distribution of PLV values over the parieto-occipital (PO) brain regions and the topographically thresholded functional connections are shown in Fig.4. The threshold was estimated as the mean+S.D. from the PO sub-graph from level 1. PLV values were higher in level 1 compared to level 5 as can be seen in Fig.4 (c) & (d).

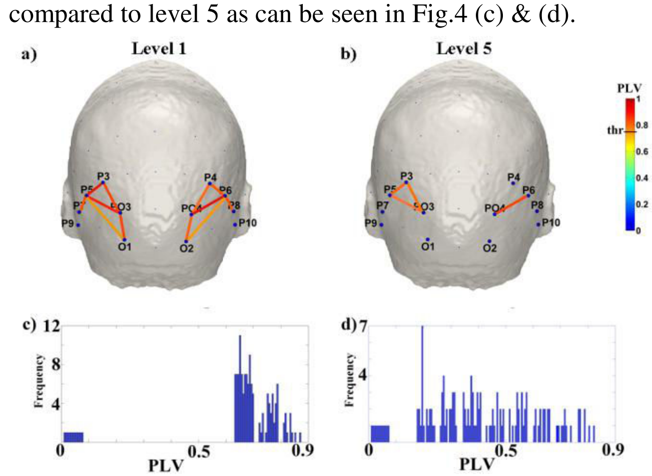

Fig.4. Topography of Parieto-occipital functional connections in level 1 (a) and level 5 (b) after thresholding; c) and d) indicate the distribution of PLV values averaged across trials from PO brain regions in level 1 and 5 correspondingly. The threshold level is indicated in the color bar.

IV. RESULTS

A. Machine learning validation

The TSA algorithm, followed by a k-nearest-neighbor classifier (with k=3), was tested on trial-based connectivity data from all bands. The following results have been obtained through a cross-validation scheme that shuffle the trials and get 90% for training and 10% for testing. The following table summarizes the average classification rates derived after applying the above cross-validation scheme 100 times. Our analysis was based on both neighbors of PO7 and PO8 and also to each one separately. Bilaterally approach includes also inter-hemispheric connections. The selected options for TSA were: *Weight mode*=Heat Kernel; *Neighbor mode*=1; Supervised learning; *Number of dimensions*=3. Table 1 summarizes the classification performance of TSA+knn and of LDA+knn (Linear Discriminant Analysis) as a baseline validation feature extraction algorithm.

TABLE I
SUMMARY OF THE CLASSIFICATION PERFORMANCE

	Bilaterally PO		Left PO		Right PO	
	TSA	LDA	TSA	LDA	TSA	LDA
δ	93.59	83.28	87.81	80.39	91.68	79.31
θ	96.70	85.65	89.86	81.29	90.40	77.41
α_1	96.13	86.42	91.13	82.14	90.27	80.28
α_2	95.00	84.12	90.31	79.41	87.63	79.31
β_1	94.40	83.49	87.63	78.39	86.54	76.40
β_2	97.59	87.21	83.86	76.25	85.86	76.05
γ	95.86	87.34	84.86	75.83	87.54	76.91

Values are mean correct classification in percentage based on three different sub-graphs: bilateral PO sub-graph, left PO sub-graph and right PO sub-graph (PO = Parieto-occipital). TSA = tensor subspace analysis, LDA = linear discriminant analysis.

V. DISCUSSION

Based on our experiments, that tensorial treatment of FCGs increased the discriminability between easy and difficult mental workload task. Future work should test whether a more detailed assessment of the cognitive load is feasible in the derived subspace (the learning task should be switched from classification to regression). Our network analysis revealed two sensors the connectivity of which varied systematically across frequency bands based on the difficulty level of the task (PO7 & PO8). This result is in accordance with the anticipation that parietal brain regions participate in visual identification of digits and in quantity representation [15]. Even though prefrontal cortex is vital for the on-line maintenance of the intermediate results of a calculation providing a "working memory" workspace especially in the difficult addition [15], classification performance based on frontal areas in θ band were lower compared to PO brain regions (data not shown here). The investigation of the interconnectivity pattern changes between the PO regions and prefrontal cortex under different cognitive load is the subject of a future study. The introduced methodological approach can find application in many other situations where brain state has to be decoded from functional connectivity structure. After extensive validation, it might prove an extremely useful tool for on-line applications of human-machine interaction.

REFERENCES

[1] T. O. Zander, and C. Kothe, "Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general," *J. Neural. Eng.,* vol. 8, no. 2, pp. 025005, Apr, 2011.

[2] J. Richiardi, H. Eryilmaz, S. Schwartz *et al.*, "Decoding brain states from fMRI connectivity graphs," *Neuroimage,* vol. 56, no. 2, pp. 616-26, May 15, 2011.

[3] H. Shen, L. Wang, Y. Liu *et al.*, "Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fMRI," *Neuroimage,* vol. 49, no. 4, pp. 3110-21, Feb 15, 2010.

[4] L. Pollonini, U. Patidar, N. Situ *et al.*, "Functional connectivity networks in the autistic and healthy brain assessed using Granger causality," *Conf. Proc. IEEE Eng. Med. Biol. Soc.,* vol. 2010, pp. 1730-3, 2010.

[5] X. He, D. Cai, and P. Niyogi, "Tensor subspace analysis," *NIPS*, 2005.

[6] P. Niyogi, "Locality preserving projections," *NIPS 2002,* vol. 16, pp. 153-160, 2002.

[7] B. Rebsamen, K. Kwok, and T. B. Penney, "Evaluation Of Cognitive Workload From EEG During A Mental Arithmetic Task," *Proc. of Hum. Factors Ergon. Society Ann. Meeting,* vol. 55, pp. 1342-1345, 2011.

[8] S. I. Dimitriadis, N. A. Laskaris, A. Tzelepi *et al.*, "Analyzing functional brain connectivity by means of commute times: a new approach and its application to track event-related dynamics," *IEEE Trans. Biomed. Eng.,* vol. 59, no. 5, pp. 1302-9, May, 2012.

[9] A. Delorme, M. Westerfield, and S. Makeig, "Medial prefrontal theta bursts precede rapid motor responses during visual selective attention," *J. Neurosci.,* vol. 27, no. 44, pp. 11949-59, Oct 31, 2007.

[10] A. Delorme, and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods,* vol. 134, no. 1, pp. 9-21, Mar 15, 2004.

[11] S. I. Dimitriadis, K. Kassiani, N. A. Laskaris *et al.*, "Surface EEG shows that functional segregation via phase coupling contributes to the neural substrate of mental calculations," *Brain Cogn.,* vol. 80, no. 1, pp. 45-52, Oct, 2012.

[12] J. P. Lachaux, E. Rodriguez, M. L. Van Quyen *et al.*, "Studying single-trials of phase synchronous activity in the brain," *Int. J. Bifurcat. Chaos,* vol. 10, no. 10, pp. 2429-2439, Oct, 2000.

[13] S. I. Dimitriadis, N. A. Laskaris, V. Tsirka *et al.*, "Tracking brain dynamics via time-dependent network analysis," *J. Neurosci. Methods,* vol. 193, no. 1, pp. 145-155, Oct 30, 2010.

[14] S. Achard, and E. Bullmore, "Efficiency and cost of economical brain functional networks," *Plos Comput. Biol.,* vol. 3, no. 2, pp. 174-183, Feb, 2007.

[15] H. D. Shane, "The number sense: How mathematical knowledge is embedded in our brains - Dehaene,S," *Library Journal,* vol. 122, no. 16, pp. 116-116, Oct 1, 1997.