On Temporal Connectivity of PFC Via Gauss–Markov Modeling of fNIRS Signals

Sergül Aydöre*, M. Kıvanç Mıhçak, *Member, IEEE*, Koray Çiftçi, and Ata Akın, *Member, IEEE*

***Abstract*—Functional near-infrared spectroscopy (fNIRS) is an optical imaging method, which monitors the brain activation by measuring the successive changes in the concentration of oxy- and deoxyhemoglobin in real time. In this study, we present a method to investigate the functional connectivity of prefrontal cortex (PFC) Sby applying a Gauss–Markov model to fNIRS signals. The hemodynamic changes on PFC during the performance of cognitive paradigm are measured by fNIRS for 17 healthy adults. The color-word matching Stroop task is performed to activate 16 different regions of PFC. There are three different types of stimuli in this task, which can be listed as incongruent stimulus (IS), congruent stimulus (CS), and neutral stimulus (NS), respectively. We introduce a new measure, called "information transfer metric" (ITM) for each time sample. The behavior of ITMs during IS are significantly different from the ITMs during CS and NS, which is consistent with the outcome of the previous research, which concentrated on fNIRS signal analysis via color–word matching Stroop task. Our analysis shows that the functional connectivity of PFC is highly relevant with the cognitive load, i.e., functional connectivity increases with the increasing cognitive load.**

***Index Terms*—Functional connectivity, gauss markov model, near infrared spectroscopy, stroop test.**

I. Introduction

IT HAS been shown that different regions of the human brain work in a connected manner while executing a cognitive task. This aspect of the brain is related to its specialization and integration capability, i.e., particular regions of the brain are both specialized for certain tasks and share information between each other. Revealing this, connectivity structure of the brain is gaining more interest in recent years. For this purpose, dynamic causal models (DCMs) [1] and vector autoregressive models (VARs) [2] have been employed. Connectivity assumes that there is a link between one neural system and another one anatomically (anatomical connectivity), statistically (functional connectivity), and causally (effective connectivity) [3]. Anatomical connectivity necessitates physiological and structural connections in neural system. Functional connectivity is based on correlations between spatially distributed neurophysiological events, but it does not give any reference to physical connections. Granger causality employs VAR models to investigate functional connectivity [4]. Without any model, it checks for the interactions between the whole set of measurements and determines the significant ones. Effective connectivity, on the other hand, is based on DCM and begins with an underlying model, well-defined regions and prior connectivity probabilities between them. Consequently, via building a Bayesian model, it is possible to find the posterior probabilities of different models that explain the measured data, and subsequently enable choosing the best one [5]. There is a potential relationship between those connectivity types where the link might be segregated and integrated. Previous studies have shown that PFC becomes activated during high-order cognitive tasks, attention and motor tasks, and that this activation can reliably be measured with functional near-infrared spectroscopy (fNIRS) [6]–[8]. fNIRS measures oxygenated and deoxygenated hemoglobin (oxy-Hb and deoxy-Hb, respectively) within the probed regions of the brain [9]. Several signals processing techniques have been proposed in fNIRS studies that offered improved estimation of the changes in hemodynamical concentration. These techniques ranged from bandpass filtering and then simple averaging [10], [11] to more advanced and recently general linear model approach and their statistical significance during activation [12], [13]. These studies are generally univariate and concentrate on determining the activated regions. However, whether it is possible to investigate the functional connectivity with fNIRS or not is still a challenging question to answer.

In this study, we explore the functional connectivity of the brain using fNIRS modality, by employing autoregressive (AR) models. AR model is a common way to establish a causal link (correlation) between the present and past of a process. The key assumption employed in this study is that a particular measurement at a particular instant of time may be explained by the recent past of the total set of measured values. Thus, the proposed AR model is exploited as a tool to understand the "information transfer" among different brain regions. The time-dependent AR coefficient matrices demonstrate the characteristics of the information transfer. The cognitive task employed is the color–word matching Stroop task, which is known to be a good cognitive activator of the PFC [14]. The main hypothesis in this study is that

Manuscript received September 8, 2008; revised December 16, 2008 and March 3, 2009. First published April 28, 2009; current version published February 17, 2010. This work was supported by the Boğaziçi University Research Fund under Project 04X102D and Project 04S101, and by the Turkish State Planning Organization under Project 03K120250 and Project 03K120240. The work of M. K. Mıhçak was supported in part by Turkiye Bilimsel ve Teknolojik Arastirma Kurumu (TÜBİTAK) Career Award 106E117. The work of S. Aydöre was supported in part by the Turkiye Bilimler Akademisi-Genc Bilim Insanlarini Odullendirme Programi (TÜBA-GEBIP) Award and in part by the TÜBİTAK Graduate Scholarship. *Asterisk indicates corresponding author.*

*S. Aydöre is with the Department of Electrical and Electronics Engineering, Boğaziçi University, 34342 Istanbul, Turkey (e-mail: sergul.aydore@boun.edu.tr).

M. K. Mıhçak is with the Department of Electrical and Electronics Engineering, Boğaziçi University, 34342 Istanbul, Turkey (e-mail: kivanc.mihcak@boun.edu.tr).

K. Çiftçi is with the Department of Electronics and Telecommunication Engineering, Namık Kemal University, Tekirdag, Turkey (e-mail: koray.ciftci@gmail.com).

A. Akın is with the Institute of Biomedical Engineering, Boğaziçi University, 34342 Istanbul, Turkey (e-mail: ata.akin@boun.edu.tr@boun.edu.tr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2009.2020792

"the connectivity behavior of the brain is related to the cognitive load." We conjecture that with increasing cognitive load, information transfer among the PFC regions should increase, which will exhibit itself in the derived AR coefficient matrices. Our approach can be categorized under the "functional connectivity" framework, because the model is concentrated on statistical dependencies and there is no underlying physiological model. Since the model is time-varying, a successive set of matrices are determined for a certain period of measurements. In neuroimaging, the goal is generally to find some "average" or "common" patterns for a group of subjects. For this purpose, we bring together the actual data of different subjects with the assumption that the behaviors of AR coefficient matrices that show the temporal connectivity are common for a specific group of subjects.

II. Notation

In this paper, boldface uppercase letters denote matrices and boldface lowercase letters denote vectors; corresponding nonbold letters with subscripts represent individual elements: given a vector $\mathbf{a}$, a_i denotes its ith element. Similarly, given a matrix $\mathbf{A}$, $A_{i,j}$ denotes its (i,j)th element. Given a random variable X, $p_X(x)$ denotes its probability density function; further, given a matrix $\mathbf{A}$, $\mathbf{A}^T$ denotes its transpose. We use $\|.\|$, $\exp(\cdot)$, and $\mathcal{N}(\mu, \mathbf{\Sigma})$, for the Euclidean norm, the exponential function, and the probability density function of a Gaussian random vector with mean vector μ and covariance matrix $\mathbf{\Sigma}$. The abbreviations "TCM," "ITM," and "ML" are shorthands for "temporal connectivity matrix," "information transfer metric," and "maximum likelihood," respectively.

III. Problem Formulation and Methodology

We use the quantities M, J, and K to denote the number of different regions of PFC (i.e., number of "channels" in our terminology), maximum amount of time samples, and total number of persons, respectively. Also, the corresponding indices m, j, and k denote the channel m, time sample t_j, and subject k, respectively, where $m \in \{1, 2, \ldots, M\}$, $k \in \{1, 2, \ldots, K\}$, and $j \in \{1, 2, \ldots, J\}$. The collected datum from the mth channel at time instant t_j from the kth subject is represented by $y_m^k(t_j)$. Then, for all j and k, the vector $\mathbf{y}_j^k \in \mathbb{R}^M$, where $\mathbf{y}_j^k \stackrel{\triangle}{=} [y_1^k(t_j), y_2^k(t_j), \ldots, y_M^k(t_j)]$, represents the data from kth subject at time instant t_j from all channels.

Our initial experiments revealed that the data $\{\mathbf{y}_j^k\}$ at consecutive time instances are heavily *locally* correlated. Furthermore, it is intuitively clear that, at any given time instant, all significant correlations about the causal past are mainly captured in the recent past, which, in turn, suggest the usage of a *Markovian* model to represent the system[1]. Following the conventional "AR signal modeling" approach in statistical signal processing, we utilize a Gauss–Markov AR-1 (i.e., with memory 1) model. This model is represented by

$$\mathbf{y}_{j+1}^k = \mathbf{A}_j \mathbf{y}_j^k + \mathbf{n}_j^k, \tag{1}$$

[1]Such Markovian models have traditionally been employed in statistical signal processing to model and process naturally locally correlated signals, such as digital images and speech data.

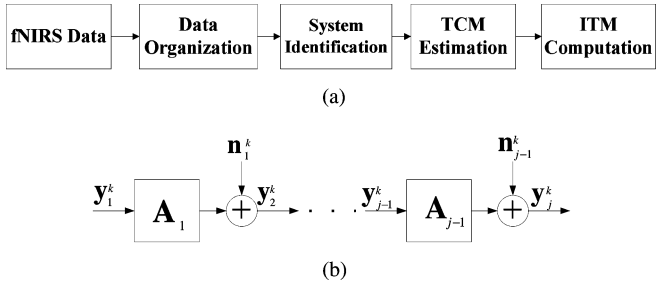

Fig. 1. General description of the proposed setup. (a) Block diagram of the system. (b) Diagram of proposed AR-1 model.

where $\mathbf{A}_j \in \mathbb{R}^{M \times M}$, and the noise vector $\mathbf{n}_j^k \in \mathbb{R}^M$. Here, $\mathbf{n}_j^k \stackrel{\triangle}{=} [n_1^k(t_j), n_2^k(t_j), \ldots, n_M^k(t_j)]$ accounts for the system model imperfections and measurement noise. In our model, the following assumptions are made.

A1: At any given time j, the matrix $\mathbf{A}_j$ is the same for all the subjects $1 \leq k \leq K$ since a common connectivity pattern is aimed to be estimated for a group of subjects.

A2: The noise vector $\mathbf{n}_j^k$ is independent identically distributed (i.i.d.) Gaussian noise with mean 0 and covariance matrix $\mathbf{\Sigma}_j^k$, for all j and k.

A3: The noise covariance matrix $\mathbf{\Sigma}_j^k$ is invariant over different subjects and time. Hence, we drop the subscript and superscript, and use $\mathbf{\Sigma}$ to denote the covariance matrix for all time $\{t_j\}_{j=1}^J$ and subjects $1 \leq k \leq K$. Also, $\mathbf{\Sigma}$ is positive-definite.

A4: The measured data $\{\mathbf{y}_j^k\}$ are statistically independent for all different subjects $1 \leq k \leq K$.

The diagram representing our approach and the diagram of the proposed AR-1 model are given in Fig. 1(a) and (b), respectively. Based on our model, we show that the structure of the estimates of the TCMs $\{\mathbf{A}_j\}$ yield highly significant and distinguishing information about the mental state of the subject. Mathematically, various estimation techniques can be employed to produce reliable estimates of $\{\mathbf{A}_j\}$. In particular, we employ an approximate ML estimation technique, which is discussed in detail in the subsequent sections. The setup that is used in this study yields approximate ML estimates of matrices $\{\mathbf{A}_j\}$, which constitute TCM estimates that (approximately) minimize the probability of estimation error under uniform prior assumption.

IV. ML Estimates of TCMs

A. Mathematical Preliminaries

In this section, we present several definitions that will be used in the derivation of the fundamental result (which is the topic of the next section).

Definition 4.1: The mapping $\mathcal{M} : \mathbb{R}^{M \times M} \rightarrow \mathbb{R}^{M^2}$ is such that, for any $\mathbf{A} \in \mathbb{R}^{M \times M}$, $\mathbf{a} \in \mathbb{R}^{M^2}$, if $\mathbf{a} = \mathcal{M}(\mathbf{A})$, we have $a_{(i-1)M+j} = A_{ij}$, where $1 \leq i, j \leq M$. In other words, the mapping $\mathcal{M}$ "reshapes" a matrix into a vector via rearranging all of its elements

Definition 4.2: The mapping $\mathcal{T}:\mathbb{R}^M \rightarrow \mathbb{R}^{M\times M^2}$ is such that, for any $\mathbf{a}\in\mathbb{R}^M$, $\mathbf{A}\in\mathbb{R}^{M\times M^2}$, if $\mathbf{A}=\mathcal{T}(\mathbf{a})$, we have

$$A_{m,n}=\begin{cases} a_{(m-1)M+n}, & \text{if } ((m-1)M+1)\leq n\leq mM \\ 0, & \text{else}\end{cases}$$

where $1\leq m\leq M$ and $1\leq n\leq M^2$. Thus, the mapping $\mathcal{T}$ forms a sparse matrix (given an input vector) via repeating the vector's elements and shifting appropriately in each row in a "staircase" fashion

Definition 4.3: The mapping $\mathcal{S}$, whose domain is K-fold cross-product of $\mathbb{R}^M$ and range is $\mathbb{R}^{KM}$, is such that, for any $\mathbf{a}^1,\mathbf{a}^2,\ldots,\mathbf{a}^K\in\mathbb{R}^M$, $\mathbf{b}\in\mathbb{R}^{KM}$, if $\mathbf{b}=\mathcal{S}(\mathbf{a}^1,\mathbf{a}^2,\ldots,\mathbf{a}^K)$, we have $b_{(k-1)M+m}=a_m^k$, where $1\leq m\leq M$, $1\leq k\leq K$. Thus, the mapping $\mathcal{S}$ concatenates K vectors each of which is of length M.

Definition 4.4: The mapping $\mathcal{L}$, whose domain is K-fold cross-product of $\mathbb{R}^{M\times M^2}$ and range is $\mathbb{R}^{KM\times M^2}$, is such that, for any $\mathbf{A}^1,\mathbf{A}^2,\ldots,\mathbf{A}^K\in\mathbb{R}^{M\times M^2}$, $\mathbf{B}\in\mathbb{R}^{KM\times M^2}$, if $\mathbf{B}=\mathcal{L}(\mathbf{A}^1,\mathbf{A}^2,\ldots,\mathbf{A}^K)$, we have $B_{(k-1)M+m,n}=A^k_{m,n}$, where $1\leq m\leq M$, $1\leq n\leq M^2$, $1\leq k\leq K$. Thus, the mapping $\mathcal{L}$ concatenates K matrices each of which is of size $M\times M^2$

Note that these mappings are clearly all invertible.

B. Main Result

In this section, we derive the ML estimate $\hat{\mathbf{A}}_{j,\mathrm{ML}}\in\mathbb{R}^{M\times M}$ of $\mathbf{A}_j$ for all j (cf., (1) via employing the model introduced in Section III, and subsequently discuss how this result can be used in computing TCMs. We proceed with providing the definition of the ML estimate $\hat{\mathbf{A}}_{j,\mathrm{ML}}$.

Definition 4.5: For all $j\in\{1,2,\ldots,J\}$, the ML estimate, $\hat{\mathbf{A}}_{j,\mathrm{ML}}$, of $\mathbf{A}_j$ is given by

$$\hat{\mathbf{A}}_{j,\mathrm{ML}}\stackrel{\triangle}{=}\underset{\mathbf{A}_j\in\mathbb{R}^{M\times M}}{\operatorname{argmax}}\; p_{\mathbf{y}}\left(\{\mathbf{y}_{j+1}^l\}_{l=1}^K\middle|\{\mathbf{y}_j^l\}_{l=1}^K,\mathbf{A}_j\right) \quad (2)$$

where we have

$$p_{\mathbf{y}}\left(\{\mathbf{y}_{j+1}^l\}_{l=1}^K\middle|\{\mathbf{y}_j^l\}_{l=1}^K\right)=\prod_{k=1}^K p_{\mathbf{y}}\left(\mathbf{y}_{j+1}^k\middle|\mathbf{y}_j^k,\mathbf{A}_j\right) \quad (3)$$

as result of the fact that the observations $\{\mathbf{y}_j^k\}_{j,k}$ are independent over k for all j (recall assumption A4).

Since Gaussianity is preserved under linear operators and applying this result in (1), it can be concluded that $\mathbf{y}_{j+1}^k$ is also Gaussian conditioned on $\mathbf{y}_j^k$ and $\mathbf{A}_j$. Furthermore, a quick inspection of (1) suggests having $p_{\mathbf{y}}(\mathbf{y}_{j+1}^k|\,\mathbf{y}_j^k,\mathbf{A}_j)$ as $\mathcal{N}(\mathbf{A}_j\mathbf{y}_j^k,\mathbf{\Sigma}_j^k)$, which, in turn, reduces to $\mathcal{N}(\mathbf{A}_j\mathbf{y}_j^k,\mathbf{\Sigma})$ per assumptions A2 and A3. Thus, the probability density function $p_{\mathbf{y}}(\mathbf{y}_{j+1}^k|\,\mathbf{y}_j^k,\mathbf{A}_j)$ is $\mathcal{N}(\mathbf{A}_j\mathbf{y}_j^k,\mathbf{\Sigma})$, induced by the noise term $\mathbf{n}_j^k$ for all j, k.

Proposition 4.1: The solution of the optimization problem (2) is given by $\hat{\mathbf{A}}_{j,\mathrm{ML}}=\mathcal{M}^{-1}(\hat{\mathbf{a}}_{j,\mathrm{ML}})$, for all j, where

$$\hat{\mathbf{a}}_{j,\mathrm{ML}}\stackrel{\triangle}{=}\min_{\mathbf{a}_j\in\mathbb{R}^{M^2}}\|\bar{\mathbf{z}}_{j+1}-\bar{\mathbf{Z}}_j\mathbf{a}_j\|^2=\left(\bar{\mathbf{Z}}_j^T\bar{\mathbf{Z}}_j\right)^{-1}\bar{\mathbf{Z}}_j\bar{\mathbf{z}}_{j+1}$$

$$\bar{\mathbf{Z}}_j\stackrel{\triangle}{=}\mathcal{L}\left(\mathbf{Z}_j^1,\mathbf{Z}_j^2,\ldots,\mathbf{Z}^K\right),\quad \bar{\mathbf{z}}_j\stackrel{\triangle}{=}\mathcal{S}\left(\mathbf{z}_j^1,\mathbf{z}_j^2,\ldots,\mathbf{z}^K\right)$$

$$\mathbf{Z}_j^k\stackrel{\triangle}{=}\mathbf{\Sigma}^{-1/2}\mathbf{Y}_j^k,\quad \mathbf{Y}_j^k\stackrel{\triangle}{=}\mathcal{T}\left(\mathbf{y}_j^k\right),\quad \mathbf{z}_j^k\stackrel{\triangle}{=}\mathbf{\Sigma}^{-1/2}\mathbf{y}_j^k \quad (4)$$

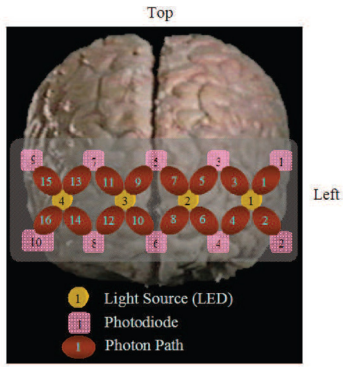

Fig. 2. Approximate placement of the fNIRS head probe on the PFC.

for all $j\in\{1,2,\ldots J\}$, $k\in\{1,2,\ldots,K\}$, assuming that $\{\bar{\mathbf{Z}}_j\}_{j=1}^J$ are all full-rank matrices.

Proof: See the Appendix. ■

The result (4) is derived based on the assumption that the noise covariance matrix $\mathbf{\Sigma}$ is known. However, in practice, $\mathbf{\Sigma}$ is unknown; as such, we proceed with using the empirical estimate $\hat{\mathbf{\Sigma}}$ instead of $\mathbf{\Sigma}$ in (4) to reach practically useful results: $\hat{\mathbf{\Sigma}}\stackrel{\triangle}{=}(1/J)(1/K)\sum_{k=1}^K\sum_{j=1}^J[(\hat{\mathbf{n}}_j^k)\cdot(\hat{\mathbf{n}}_j^k)^T]$, where $\hat{\mathbf{n}}_j^k\stackrel{\triangle}{=}\mathbf{y}_j^k-(1/K)\sum_{l=1}^K\mathbf{y}_j^l$. Because of the aforementioned reason and the resulting approach, we term the resulting estimate of $\mathbf{A}_j$ as *approximate ML estimate*. The experimental results provided in Section V are computed based on this approach. Note that, the result (4) assumes that $\{\bar{\mathbf{Z}}_j\}_{j=1}^J$ are full-rank for all j (which ensures the invertibility of $\bar{\mathbf{Z}}_j^T\bar{\mathbf{Z}}_j$ for all j). In our experiments (see Section V), we observed that this is indeed always the case. However, technically this does not necessarily need to hold; in such cases, a solution to the optimization problem (4) may be found, for instance, via following the MNLS (minimum norm least squares) approach, which can be expressed in terms of the singular value decomposition (SVD) of the matrix $\bar{\mathbf{Z}}_j$ for all j.

V. Experimental Studies

A. Experimental Setup

A continuous-wave near-infrared spectroscopy device (NIROXCOPE 301) that was built in Biophotonics Laboratory of Bogazici University is used during the experiments [6], [7]. This fNIRS device can sample 16 different channels on PFC simultaneously using 10 detectors and 4 LEDs. Thus, oxy-Hb data of 16 different regions on PFC for each time series can be measured. The light that the device transmits is at two wavelengths, which are 730 and 850 nm. The sampling rate of the device is 1.7 Hz. The regions of PFC where fNIRS measures are depicted in Fig. 2.

B. Experimental Protocol

The color–word matching Stroop task is used for cognitive task. Three different types of stimulus are generated with this task, namely neutral stimulus (NS), congruent stimulus (CS), and incongruent stimulus (IS). For each type of trial, two rows

TABLE I
STATISTICAL PROPERTIES

Stimulus Type	Reaction Times (sec)		Error Rates (%)	ITM Values	
	Mean	Variance	Mean	Mean	Variance
NS	0.9687	0.0037	1.3725	0.7111	0.0598
CS	1.0658	0.0117	2.156	0.7237	0.0520
IS	1.1592	0.0099	6.4706	0.8216	0.0270

of letters are displayed on the screen. The bottom row is the color of letters displayed in the top row. The subjects decide whether or not the bottom row is the true color of the word appeared in the top row. During NS trials, the word "XXXX" appears in the color red, blue, yellow, and green in top row, and the words "RED," "BLUE," "YELLOW," and "GREEN" appears in black in the bottom row. During CS (respectively IS) trials, the words "RED," "BLUE," "YELLOW," and "GREEN" appear in top row in congruent (respectively incongruent) color. It is known that during IS trials the brain activation increases during color–word Stroop task [14]. In addition, the questions asked to subjects are more difficult during IS compared with NS and CS. Therefore RTs and error rates of IS are expected to be higher than the NS and CS.

There are 30 questions for each type of stimulus and there are 5 blocks for each stimulus type and the time gap between blocks are considered as rest periods. The trials were presented in a semiblocked manner. Each block consisted of six trials. Interstimulus interval within the block was 4.5 s and the blocks were placed 20 s apart in time. The trial type within a block was homogeneous (but the arrangements of false and correct trials were altering). There were ten blocks of each type. Those blocks were presented in a random fashion. Experiments were performed in a silent, lightly dimmed room. Words were presented via an LCD screen that was 0.5 m away from the subjects. The task protocol is approved by the Ethics Review Board of Bogazici University. In this paper, oxy-Hb change are used to analyze connectivity, because the device we used (NIROX-COPE 301) has better capability of measuring oxy-Hb rather then deoxy-Hb [7]. The measured data are separated according to each stimulus type and sorted with respect to the time.

C. Experimental Results

1) Behavioral Results: We analyzed the reaction times (RTs) in seconds and percentage error rates (ERs) from data of 17 subjects (10 women, 7 men, aged 20–24 years), all college students. The statistical properties of RT and ER are given in Table I, from which we can conclude that the behavioral results are consistent with previous work on color–word Stroop task [14]. In order to compare each pair of stimulus, we apply a two-tailed pair t-test for RTs.

The RT for IS and NS ($p = 0.0011$), and RT of IS and CS are significantly different ($p = 0.0121$). However, there is no significant difference between RT of NS and CS ($p = 0.6224$), which is an expectable result because it is still a controversial topic that questions which stimulus type is more difficult [15].

2) Temporal Connectivity Analysis: We give the results for TCM that was explained in Section III mathematically. For each stimulus type, three samples from median time of corresponding block are given in Fig. 3. These matrices represents temporal connectivity between 16 different regions of PFC. The background of the matrices are 0, because we reduce the matrices by taking maximum value (i.e., maximum connectivity) for each column and row and fixed the rest to 0. Let x-axis and y-axis denote horizontal values and vertical values on a given TCM, respectively. The nonzero little squares represent the connectivity between corresponding coordinates of the matrices, i.e., channels of the fNIRS device. For instance, the square with coordinates (2,7) represents the flow of information from the seventh channel to the second channel for a specific time.

In Fig. 3, we present the instantaneous TCMs for each type of stimulus estimated over all the subjects. It is clear that temporal TCMs of NS and CS are more close to the subspace of diagonal matrices. Closeness to the subspace of diagonal matrices can be interpreted as possessing less connectivity between different regions of PFC. However, TCMs of IS are not as close as to the subspace of diagonal matrices as TCMs of NS and CS. Note that these matrices are snapshots of the TCM but not grand averages. The similarity in some of the matrices implies an increased connectivity even during a low load task (NS) but does not assure its continuity throughout the block. It is possible to calculate an averaged TCM for each stimulus type to better elucidate the nature of the connectivity pattern. Our results show that increase in cognitive load results in an increase in connectivity between PFC regions. In order to parameterize the connectivity, we define a constant α_j for each time t_j, such that $\alpha_j \stackrel{\triangle}{=} \|\mathbf{A}_j - \mathbf{A}_{j,\mathrm{diag}}\|/\|\mathbf{A}_j\|$, where $\mathbf{A}_{j,\mathrm{diag}}$ is the projection of $\mathbf{A}_j$ to the subspace spanned by all diagonal matrices. Hence, $\{\alpha_j\}$ represent the "diagonality degree" of corresponding TCM $\{\mathbf{A}_j\}$. Then, we call this metric as ITM. By definition, these metrics lie between 0 and 1 where 0 (respectively 1) describes that the matrix is very close to (respectively far from) the diagonal matrices subspace. So, if the metric is close to 0 (respectively 1), we can say that there is low (respectively high) connectivity between the regions of PFC. The statistical properties of these metrics for each type of stimulus are given in Table I, from which, we can conclude that during IS trials, temporal connectivity between regions of PFC increases that is consistent with behavioral results. In order to compare ITM values according to stimulus type, we applied two-tailed t-test. The differences between ITM values of NS and IS, and CS and IS are statistically significant ($p < 0.0001$). However, the ITM values between CS and NS did not show statistical significance ($p = 0.5312$) similar to the behavioral results. Behavioral results and temporal connectivity analysis according to stimulus type are shown simultaneously in Fig. 4. The most connected regions on PFC can be estimated by taking the mean of TCMs over the whole time series for each stimulus type. The mean values of TCMs for each stimulus are calculated and their reduction forms are shown in Fig. 5. From the figures, most connected channels are 2 and 11 for NS and CS, and 2 and 13 for IS. As shown in Fig. 2, channel 2 is located on the left side of PFC, and channels 11 and 13 are located on the right side of PFC. Hence, the result can be interpreted as follows: there is a high

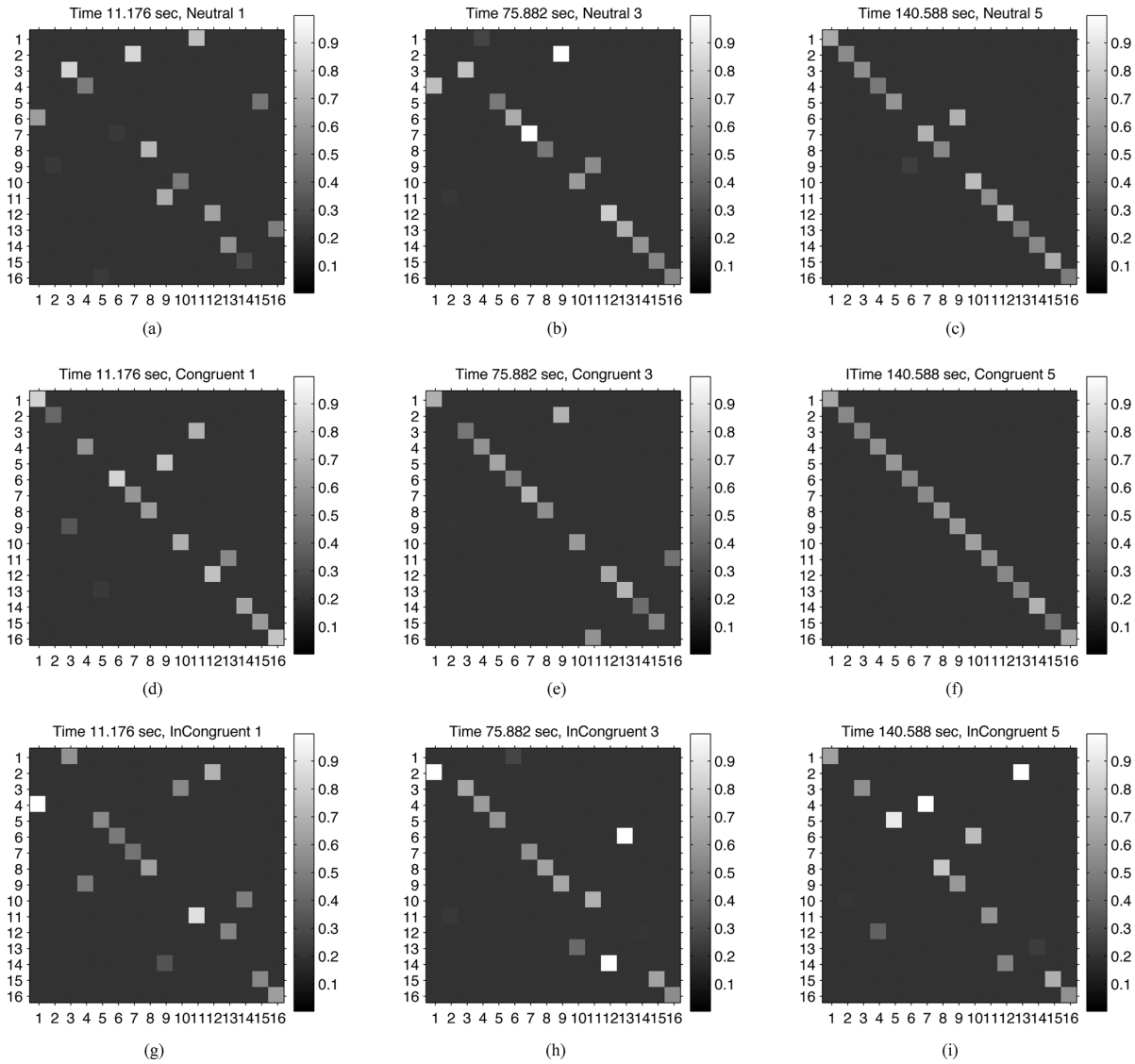

Fig. 3. Estimated temporal connectivity matrices for: (a) NS at given time within the first block, (b) NS at given time within the third block, (c) NS at given time within the fifth block, (d) CS at given time within the first block, (e) CS at given time within the third block, (f) CS at given time within the fifth block, (g) IS at given time within the first block, (h) IS at given time within the third block, and (i) IS at given time within the fifth block.

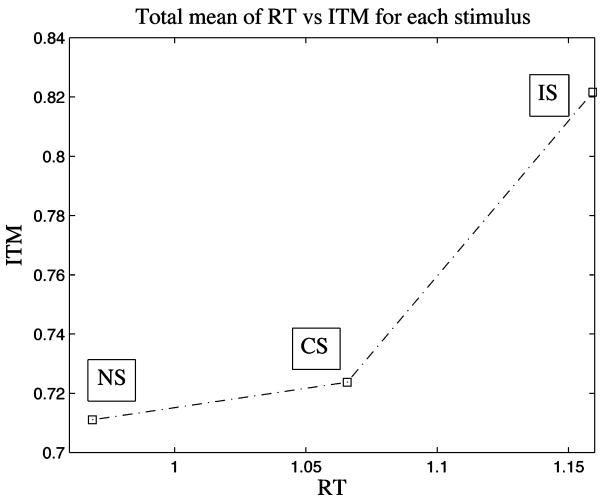

Fig. 4. Comparison of RT and ITM.

connectivity between left and right dorsolateral parts of PFC while performing cognitive paradigm. It should be noted that we have not quantified the temporal evolution of the TCMs before and after each stimulus and during a block of stimuli. The stimuli were presented at every 4.5 s, that leaves very little time for the brain hemodynamic response to evolve to a peak. Hence, the hemodynamic responses can be thought of being convolved with each other during a block of stimuli. Thus, we were under the assumption that there will be minimal change between adjacent TCMs. Besides, the working hypothesis of the study was "there are significant connectivity pattern differences between stimulus types." Hence, we decided to present an averaged metric to quantify the degree of connectivity rather than observing the evolution of this pattern around an interval of a stimulus and even during the block. All in all, we believe that the evolution of the TCM throughout a task might possess clinical significance, which is worth investigating.

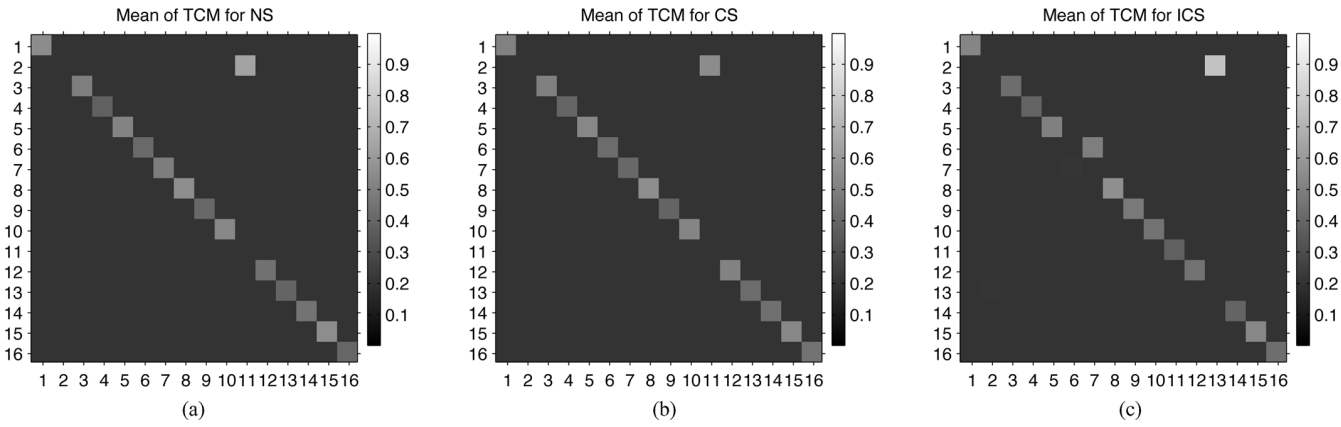

Fig. 5. Mean of the TCM over all time for each stimulus type.

VI. Conclusion

This study proposes a novel method to identify the functional connectivity of the brain via Gauss–Markov modeling of fNIRS signals. The method is based on monitoring the information transfer among the different regions of the brain. Accordingly, this study introduces a simple methodology to be able to perform functional connectivity studies using fNIRS signals. To verify our approach, we implemented the method on fNIRS signals measured during a Stroop test.

The results corroborate the conjecture that the structure of the ITM depends on the cognitive load of the task. TCM results are consistent with the hypothesis that information transfer among the regions of PFC will increase with the increasing cognitive load. For NS and CS trials, TCMs are close to subspace of diagonal matrices, whereas for IS trials, TCMs lose their diagonal structures and nondiagonal terms appear, which demonstrates that the interactions among the brain regions increase. ITM measures are defined in order to identify those TCMs. Applying two-tailed t-test to ITMs for each stimulus type shows the significant difference between ITMs during IS, and ITMs during NS and CS. Our results are consistent with the literature, because recent functional magnetic resonance imaging (fMRI) investigations have found that the lateral prefrontal cortex (PFC) areas are preferentially activated to increasing demand [16], [17]. Stroop task is known to induce bilateral activations of the dorsolateral PFC (DLPFC). In our results, we also found the highest connections between the second and eleventh detectors for the NS and CS cases (second and thirteenth for the IS case) where the second one lies on the far left dorsolateral PFC and eleventh (and thirteenth) is approximately 4.5 cm (and 3 cm) to the right from the midline on the forehead, respectively. Our behavioral data are also consistent with the literature whereby we see strong interference between the neutral and incongruent stimuli, as observed by others [15]. Hence, we conclude that the Stroop activity determined by our connectivity analysis shows a bilateral character. Schroeter *et al.* [14] have shown that the activation is bilateral on the PFC for oxy-Hb, deoxy-Hb, and total Hb. However, Ehlis *et al.* [13] and Ciftci *et al.* [7] have found that the activation is left lateral for oxy-Hb. Although all of these works focus on activation instead of connectivity, our results coincide with Schroeter *et al*.

TCMs are calculated for each instant of time. For all of the three stimulus types, we observe that information transfer characterized by the TCM vary with time. However, there is not a well-defined relationship between ITM and time. A hypothesis might be that as the trial block progresses, the ITM metric should decrease because of learning, habituation, or getting tired. The fact that this hypothesis was not proved by the experimental results leads us to the conclusion that ITM truly reflects the cognitive load without being affected much from secondary factors. The relation between the behavioral results of the Stroop test and functional connectivity has been also investigated by a PET study, where the researchers confirmed that the canonical variants (a measure of connectivity) increases from congruent stimuli to incongruent stimuli [18]. Similar results were found by Strother *et al.* [19]. In a recent work, Szatkowska *et al.* [20] have observed that left DLPFC showed a significant condition dependent change in the strength of influence conveyed through the pathway during a verbal two back working memory task under two reinforcement conditions.

It may be argued that the proposed approach is sensitive to sampling rate of the existing fNIRS device. The main spectral component of the hemodynamics signals lie around 0.1 Hz [21]–[23]. Although some researchers argue that there might be some fast optical signals with a higher bandwidth, it is still a controversial topic in fNIRS community [24]. Hence, according to the Nyquist criterion, sampling rate of 1.7 Hz is sufficient to quantify cognition-related hemodynamic changes. Higher sampling rate may lead to calculation of ITM metric of the very close time samples that are potentially similar. On the other hand, low sampling rate may cause information loss. Yet, the detailed sensitivity analysis of the method may be a further research topic.

An important drawback of fNIRS studies is that the measurements are dependent on the differential path length factor (DPF) [9]. It was shown that DPF may vary between subjects mainly depending on age [25]. Since it is hard to measure it exactly, DPF is generally assumed to be constant between subjects. An important advantage of the metric (ITM) that we proposed in this study is its independence of DPF.

The proposed method in this study may complement connectivity analysis on PFC during cognitive tasks. Our approach

is distinguished from the well-known methods of Granger test, VAR, and DCM at some points. DCM is a method of finding effective connectivity that treats the brain as a deterministic nonlinear system [1], while the method we have proposed estimates the functional connectivity with the assumption that the model is linear and stochastic. Granger test is a method employed to extract dependencies between pairs of time series and maybe misleading when there is a relationship between three or more variables. VAR is the generalization of the Granger test that is used when there are multidimensional time series. Both methods assume the data to be stationary and measure one dependency measure from the whole dataset. However, we do not use the assumption of stationarity. We calculate the dependency between 16 different regions at each time instance by utilizing 17 different subjects' data.

When we evaluate the behavioral results of the Stroop test and our results concurrently, we conclude that the proposed method is appropriate to be used in the analysis of fNIRS signals. The method may lead to new findings in the connectivity analysis of fNIRS signals. In the future, this method might be used for elucidating the pathophysiology of mental diseases as well as modeling the development of cognition.

APPENDIX

PROOF OF PROPOSITION IV.1

By the definition of ML estimate (see Definition IV-B) and using (3) in (2), we get

$$\hat{\mathbf{A}}_{j,\mathrm{ML}} = \underset{\mathbf{A}_j \in \mathbb{R}^{M\times M}}{\operatorname{argmax}} \sum_{k=1}^{K} \log p_{\mathbf{y}}\left(\mathbf{y}_{j+1}^{k} \middle| \mathbf{y}_{j}^{k}, \mathbf{A}_j\right) \tag{A1}$$

where (A1) follows from the monotonicity of the $\log(\cdot)$ function. Next, we note that, conditioned on $\mathbf{y}_j^k$ and $\mathbf{A}_j$, $\mathbf{y}_{j+1}^k \sim \mathcal{N}(\mathbf{A}_j\mathbf{y}_j^k, \boldsymbol{\Sigma})$ for all j, k. Hence, we get

$$\hat{\mathbf{A}}_{j,\mathrm{ML}} = \underset{\mathbf{A}_j \in \mathbb{R}^{M\times M}}{\operatorname{argmin}} \sum_{k=1}^{K} \|\boldsymbol{\Sigma}^{-1/2}\left(\mathbf{y}_{j+1}^{k} - \mathbf{A}_j\mathbf{y}_j^{k}\right)\|^2 \tag{A2}$$

where (A2) follows from the fact that $\boldsymbol{\Sigma}$ is positive-definite per assumption. Next, recalling the definitions $\mathbf{a}_j = \mathcal{M}(\mathbf{A}_j)$, and $\mathbf{Y}_j^k = \mathcal{T}(\mathbf{y}_j^k)$, we observe that the problem of $\hat{\mathbf{a}}_{j,\mathrm{ML}} = \arg\min_{\mathbf{a}_j \in \mathbb{R}^{M^2}} \sum_{k=1}^{K} \|\boldsymbol{\Sigma}^{-1/2}(\mathbf{y}_{j+1}^k - \mathbf{Y}_j^k\mathbf{a}_j^k)\|^2$, is "equivalent" to the problem (A2), i.e., there is a one-to-one mapping between $\hat{\mathbf{a}}_{j,\mathrm{ML}}$ and $\hat{\mathbf{A}}_{j,\mathrm{ML}}$, given by $\hat{\mathbf{A}}_{j,\mathrm{ML}} = \mathcal{M}^{-1}(\mathbf{a}_{j,\mathrm{ML}})$ due to the construction of the mappings $\mathcal{M}(\cdot)$ and $\mathcal{T}(\cdot)$. Next

$$\hat{\mathbf{a}}_{j,\mathrm{ML}} = \underset{\mathbf{a}_j \in \mathbb{R}^{M^2}}{\operatorname{argmin}} \sum_{k=1}^{K} \|\mathbf{z}_{j+1}^{k} - \mathbf{Z}_j^{k}\mathbf{a}_j^{k}\|^2 \tag{A3}$$

$$= \underset{\mathbf{a}_j \in \mathbb{R}^{M^2}}{\operatorname{argmin}} \|\bar{\mathbf{z}}_{j+1} - \bar{\mathbf{Z}}_j\mathbf{a}_j\|^2 \tag{A4}$$

$$= \underset{\mathbf{a}_j \in \mathbb{R}^{M^2}}{\operatorname{argmin}} \left(\bar{\mathbf{Z}}_j^T\bar{\mathbf{Z}}_j\right)^{-1}\bar{\mathbf{Z}}_j\bar{\mathbf{z}}_{j+1} \tag{A5}$$

where (A3) follows from the definitions of $\{\mathbf{z}_j^k\}$ and $\{\mathbf{Z}_j^k\}$, (A4) follows from the definitions of $\bar{\mathbf{z}}_{j+1}$, and $\bar{\mathbf{Z}}_j$, (A5) follows from the assumption that $\{\bar{\mathbf{Z}}_j\}$ are full-rank and applying the standard LS solution. Hence, the proof.

ACKNOWLEDGMENT

The authors thank the reviewers for their insightful and valuable comments that helped to improve the paper. Dr. A. Akin also acknowledges the support from Hemosoft, Inc., in the maintenance of NIROXCOPE 301.

REFERENCES

[1] K. J. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *Neuroimage*, vol. 19, pp. 1273–1302, 2003.

[2] R. Goebela, A. Roebroecka, D. Kimb, and E. Formisanoa, "Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping," *Magn. Reson. Imag.*, vol. 21, pp. 1251–1261, 2003.

[3] K. Friston, "Functional and effective connectivity in neuroimaging: A synthesis," *Human Brain Mapping*, vol. 2, pp. 56–78, 1994.

[4] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, pp. 424–438, 1969.

[5] W. D. Penny, K. E. Stephan, A. Mechelli, and K. J. Friston, "Comparing dynamic causal models," *NeuroImage*, vol. 22, no. 3, pp. 1157–1172, 2004.

[6] A. Akın and D. Bilensoy, "Cerebrovascular reactivity to hypercapnia in migraine patients measured with near-infrared spectroscopy," *Brain Res.*, vol. 1107, pp. 206–214, 2006.

[7] K. Ciftci, B. Sankur, Y. P. Kahya, and A. Akin, "Multilevel statistical inference from functional near infrared spectroscopy data during Stroop interference," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 9, pp. 2212–2220, Sep. 2008.

[8] A. Villringer, J. Planck, C. Hock, L. Scheleinkofer, and U. Dimagl, "Near infrared spectroscopy (NIRS): A new tool to study hemodynamic changes during activation of brain function in human adults," *Neurosci. Lett.*, vol. 154, pp. 101–104, 1993.

[9] A. Villringer and B. Chance, "Non-invasive optical spectroscopy and imaging of human brain function," *Trends Neurosci.*, vol. 20, pp. 4435–4442, 1997.

[10] J. H. Meek, C. E. Elwell, M. J. Khan, J. Romaya, J. S. Wyatt, D. T. Delpy, and S. Zeki, "Regional changes in cerebral haemodynamics as a result of visual stimulus measured by near infrared spectroscopy," *Proc. Biol. Sci.*, vol. 261, no. 1362, pp. 351–356, Sep. 22, 1995.

[11] J. Fuster, M. Guiou, A. Andestani, A. Cannestra, S. Sheth, Y. Zhou, A. Toga, and M. Bodner, "Near-infrared spectroscopy (NIRS) in cognitive neuroscience of the primate brain," *NeuroImage*, vol. 26, pp. 215–220, 2005.

[12] M. L. Schroeter, M. M. Bücheler, K. Miller, K. Uludag, H. Obrig, G. Lohmann, M. Tittgemeyer, A. Villringer, and D. Y. von Cramon, "Toward a standard analysis for functional near-infrared imaging," *NeuroImage*, vol. 21, pp. 283–290, 2004.

[13] A.-C. Ehlis, M. J. Herrmann, A. Wagener, and A. J. Fallgatter, "Multichannel near-infrared spectroscopy detects specific inferior-frontal activation during incongruent Stroop trials," *Biol. Psychol.*, vol. 69, pp. 315–332, 2005.

[14] M. L. Schroeter, S. Zysset, T. Kupka, F. Kruggel, and D. Y. von Cramon, "Near Infrared spectroscopy can detect brain activity during a color–word matching Stroop task in an event-related design," *Human Brain Mapping*, vol. 17, pp. 61–71, 2002.

[15] S. Zysset, K. Müller, G. Lohmann, and D. Y. von Cramon, "Color–word matching Stroop task: Separating interference and response conflict," *NeuroImage*, vol. 13, pp. 29–36, 2001.

[16] R. L. Mitchell, "The BOLD response during Stroop task-like inhibition paradigms: Effects of task difficulty and task-relevant modality," *Brain Cogn.*, vol. 59, no. 1, pp. 23–27, 2005.

[17] G. B. Kaplan, N. S. Sengor, H. Gurvit, and C. Guzelis, "Modelling the Stroop effect: A connectionist approach," *Neurocomputing*, vol. 70, no. 7–9, pp. 1414–1423, 2007.

[18] B. J. Harrison, M. Shaw, M. Yucel, R. Purcell, W. J. Brewer, S. C. Strother, G. F. Egan, J. S. Olver, P. J. Nathan, and C. Pantelis, "Functional connectivity during Stroop task performance," *NeuroImage*, vol. 24, no. 1, pp. 181–191, 2005.

[19] S. C. Strother, J. Anderson, L. K. Hansen, U. Kjems, R. Kustra, J. Sidtis, S. Frutiger, S. Muley, S. LaConte, and D. Rottenberg, "The quantitative evaluation of functional neuroimaging experiments, the NPAIRS data analysis framework," *NeuroImage*, vol. 15, pp. 747–771, 2002.

[20] I. Szatkowska, P. Bogorodzki, T. Wolak, A. Marchewka, and W. Szeszkowski, "The effect of motivation on working memory: An fMRI and SEM study," *Neurobiol. Learn. Mem.*, vol. 90, no. 2, pp. 475–478, 2008.

[21] H. Obrig, M. Neufman, R. Wenzel, M. Kohl, J. Steinbrinlk, K. Einhaupl, and A. Villinger, "Spontaneous low frequency oscillations of cerebral hemodynamics and metabolism in human adults," *Neuroimage*, vol. 12, no. 6, pp. 623–639, 2000.

[22] P. Wobst, R. Wenzel, M. Kohl, H. Obrig, and A. Villringer, "Linear aspects of changes in deoxygenated hemoglobin concentration and cytochrome oxidase oxidation during brain activation," *NeuroImage*, vol. 13, pp. 520–530, 2001.

[23] V. Toronov, M. A. Franceschini, M. Filiaci, S. Fantini, M. Wolf, A. Michalos, and E. Gratton, "Near-infrared study of fluctuations in cerebral hemodynamics during rest and motor stimulation: Temporal analysis and spatial mapping," *Med. Phys.*, vol. 24, no. 4, pp. 801–815, 2000.

[24] H. Radhakrishnan, W. Vanduffel, H. P. Deng, L. Ekstrom, D. A. Boas, and M. A. Franceschini, "Fast optical signal not detected in awake behaving monkeys," *Neuroimage*, vol. 45, no. 2, pp. 410–419, 2009.

[25] A. Duncan, J. H. Meek, M. Clemence, C. E. Elwell, L. Tyszczuk, M. Cope, and D. T. Delpy, "Optical pathlength measurements on adult head, calf and forearm and the head of the newborn infant using phase resolved optical spectroscopy," *Phys. Med. Biol.*, vol. 40, no. 2, pp. 295–304, 1995.

Authors' photographs and biographies not available at the time of publication.