Partial correlation-based functional connectivity analysis for functional near-infrared spectroscopy signals

Ata Akın

Partial correlation-based functional connectivity analysis for functional near-infrared spectroscopy signals

Ata Akin*
Acibadem University, Department of Medical Engineering, Atasehir, Istanbul, Turkey

Abstract. A theoretical framework, a partial correlation-based functional connectivity (PC-FC) analysis to functional near-infrared spectroscopy (fNIRS) data, is proposed. This is based on generating a common background signal from a high passed version of fNIRS data averaged over all channels as the regressor in computing the PC between pairs of channels. This approach has been employed to real data collected during a Stroop task. The results show a strong significance in the global efficiency (GE) metric computed by the PC-FC analysis for neutral, congruent, and incongruent stimuli (NS, CS, ICs; GE\(_N\) = 0.10 ± 0.009, GE\(_C\) = 0.11 ± 0.01, GE\(_I\) = 0.13 ± 0.015, \(p = 0.0073\)). A positive correlation (\(r = 0.729\) and \(p = 0.0259\)) is observed between the interference of reaction times (incongruent–neutral) and interference of GE values (GE\(_C\) – GE\(_N\)) computed from [HbO] signals. © 2017 SPIE Society of Photo-Optical Instrumentation Engineers (SPIE)

Keywords: functional near-infrared spectroscopy; partial correlation; functional connectivity; global efficiency; Stroop task.

Paper 170160RR received Mar. 14, 2017; accepted for publication Nov. 20, 2017; published online Dec. 14, 2017.

1 Introduction

The main challenge of the state-of-art of functional near-infrared spectroscopy (fNIRS) systems has been to accurately recover the deep cortical signals that are undoubtedly buried inside the upper layer hemodynamic signals. Since the measurements are performed from the skin surface, the photons travelling from a source to a detector will be modulated by the hemodynamic changes occurring at each layer. fNIRS techniques have been shown to suffer from the corruption of upper layer hemodynamic activity and only a fraction of the signal measured is attributed to cortical signals.1 The source–detector configuration necessary to ensure a penetration to the cortical level results in unfortunate sampling from the skin, where systemic physiological changes are reflected as hemodynamic changes. Hence, any further analysis on these signals will suffer from the corruption of these systemic fluctuations.2–10 One major analysis, where this corruption will lead to false positives, is the functional connectivity (FC) studies.

FC can be computed most easily by the use of a correlation coefficient between channels of fNIRS measurements. For a multichannel system, the correlation coefficient \(\hat{\rho}_{ij}\) between any two channels \((x_i, x_j)\) can be computed by

\[
\hat{\rho}_{ij} = \frac{C(x_i, x_j)}{\sigma_i \sigma_j},
\]

where \(C(x_i, x_j)\) is the covariance between the two channels, \(\sigma_i\) and \(\sigma_j\) are the standard deviations of the variables \(x_i\) and \(x_j\).

Due to physics of photon propagation in turbid media, the collected signal will be carrying information from the underlying tissues. Hence the signal, \(x_t\), will include cortical as well as non-cerebral tissue dynamics. The collected signal can then be modeled as a weighted sum of the activities of the underlying tissues:

\[
x_t(t) = w_b s_b^b (t) + w_c s_c^c (t) + n(t),
\]

where \(w_b\) and \(w_c\) are the contributions (i.e., percentage) of a signal [i.e., the regional brain signal, \(s_b^b\), the systemic signal, \(s_c^c\)] to the signal at the channel, and \(n(t)\) is the instrumentation and other random noises modeled as a Gaussian zero mean white process. Hence, by substituting Eq. (2) into Eq. (1), combining \(s^m = s^c (t) + n(t)\), and expanding the nominator and denominator terms, we get

\[
\hat{\rho}_{ij} = \frac{C(s_b^b, s_c^c) + C(s_b^b, s^m) + C(s_c^c, s^m) + C(s^m, s^m)}{\sqrt{\sigma^2(s_b^b) + \sigma^2(s_c^c) + 2C(s_b^b, s_c^c) + \sigma^2(s^m) + 2C(s_b^b, s^m) + 2C(s_c^c, s^m) + 2C(s^m, s^m))},
\]

but we are interested only in finding the correlation between the cortical regions:

\[
\rho_{ij} = \frac{C(s_b^b, s_b^b)}{\sigma_i \sigma_j}.
\]

*Address all correspondence to: Ata Akin, E-mail: ata.akin@acibadem.edu.tr

Journal of Biomedical Optics 22(12), 126003 (December 2017)
Since we do not have access to the individual covariances and variances in Eq. (3), we need to propose a method to extract this information. The issue, then, becomes the proper and accurate way of regressing out the contaminating signals to estimate the correlation of the brain regions, as the definition of FC requires. Several researchers have proposed to use a second detector placed closer to the source (short separation detector) that will be sensitive to changes only on the superficial layers to regress out the systemic fluctuations contaminating a detector placed farther away from the source.\textsuperscript{2,6-15} The use of an SSD has been investigated deeply due to its ambiguity of its placement (how short is short enough?).\textsuperscript{11} It also increases instrumentation and data analysis complexity, not to mention the extra burden it brings to ergonomics of the probe. Hence, many have promoted the use of advanced signal processing techniques to overcome the contamination problem. This paper is yet another proposal that advocates the use of such a technique to maintain the view that “few is better.”

We propose the use of a partial correlation (PC)-based FC analysis that removes the unwanted correlation due to extracerebral contamination. In many of fMRI and fNIRS studies, PC is considered as a suitable criterion for investigation and understanding of brain FC.\textsuperscript{16-23} FC studies using fNIRS have been conducted recently\textsuperscript{17,24-31} and have further been linked to graph theoretical approaches.\textsuperscript{32} Similarly, FC analysis with the PC approach has been proposed both by our group and others.\textsuperscript{17,24,31,34} Our group has shown that using the remaining 14 channels of a 16-channel fNIRS measurement system as the regressors of the PC analysis, we can provide reliable estimate of the FC analysis. The assumption behind this approach was that there are local systemic fluctuations at each measurement site (channel) and they are also picked up by any channel. Hence, when computing the correlation between two channels, one must take into account the distorting effect of the remaining 14 channels. This paper challenges this assumption by the fact that there is one unique systemic fluctuation embedded in each channel, and that its frequency content is different from the hemodynamic response to cognitive stimulus. This systemic fluctuation can be recovered by taking an average of the high pass filtering of the signal from each channel. Hence, our aim in this paper is to propose a theoretical framework to compute the FC of fNIRS signals based on the PC approach and a signal processing algorithm that can be adopted in computing the FC for fNIRS data.

2 Methods

The first part of Sec. 2 presents a simulation study that generates synthetic data to investigate if PC-based connectivity outperforms the conventional correlation analysis for computing FC-derived metrics. The second part of the section explains how the multichannel fNIRS data have been processed to provide PC-based FC metrics.

2.1 Simulation Analysis

Studies in photon migration in tissues have shown that photons follow a somewhat banana shape trajectory between a source and detector pair. Placing the source–detector with a certain distance apart, we can probe the tissues with a particular depth sensitivity profile similar to a banana shape, as shown in Fig. 1.

A signal received by a photodiode includes the changes in the absorption within the underlying tissues. In the near-infrared light spectrum, the highest absorbers are the [Hb] and [HbO] and a spectroscopic measurement of the absorption will yield the concentration changes of these chromophores, which can be calculated from Beer–Lambert’s law, explained in detail in other studies.\textsuperscript{1,15-33} Since absorption change within a layer of tissue at a specific wavelength $\lambda$ is formulated as $\Delta \epsilon^{\lambda} = \epsilon^{\lambda}_{c} \Delta C_{l}$, where $\epsilon^{\lambda}_{c} = \left[ \epsilon^{\lambda}_{[Hb]} \epsilon^{\lambda}_{[HbO]} \right]$ and $\Delta C_{l} = (\Delta [Hb], \Delta [HbO])^T$ for fNIRS studies, we can define a signal model for the [Hb] and [HbO] concentrations for any $l$ layer in the form of $H_{b}$ and $H_{bO}$.

2.1.1 Signal model

Let us assume that the hemodynamic changes $[H_{b}^{l}(t)]$ at each layer ($l$) for given chromophore ($C$) can be modeled as a sum of independent signal activity ($S$) weighted with layer specific weights (W), as shown below:

$$h_{b}^{l}(t) = WS. \tag{5}$$

where the entries of the matrix $W$, $w_{l,k}$ are the weights of a specific hemodynamic activity $s_{k}(t)$ at that specific layer $l$ that are the entries of the $S$ matrix. For the sake of simplicity, we will assign three signal activities to be present at each layer: $s^{b}(t)$, brain hemodynamic response function (BHRF); $s^{a}(t)$, task irrelevant systemic physiological hemodynamic fluctuations; and $s_{u}(t)$, uncorrelated instrumentation noise, modeled as $N(0, \sigma_{u})$. Hence, $S = [s^{b}(t) s^{a}(t) s_{u}(t)]$. Naturally, the weight of $s^{b}(t)$ for the first and second layers will be zero (i.e., $w_{1}^{b} = w_{2}^{b} = 0$, $w_{3}^{b} = 1$) while $0.1 \leq w_{1}^{a} = w_{2}^{a} = w_{3}^{a} \leq 1$.

Traditionally, $s^{b}(t)$ and $s^{a}(t)$ are defined as follows:

$$s^{b}(t - \theta^{b}_{l}) = (t - \theta^{b}_{l}) e^{-\left[\frac{[t-\theta^{b}_{l}]}{\tau^{b}_{l}}\right]}, \tag{6}$$

$$s^{a}(t - \theta^{a}_{l}) = \left\{ \sin \left[ 2\pi f_{1} \left( t - \theta^{a}_{l} \right) \right] + \sin \left[ 2\pi f_{2} \left( t - \theta^{a}_{l} \right) \right] \right\}, \tag{7}$$

where $s^{b}(t - \theta^{b}_{l})$ is the brain BHRF modeled as a gamma function with a delay of $\theta^{b}_{l}$, a linearly increasing value for each channel ranging from 3 to 10 s to assure a variance in the correlation values between each channel;\textsuperscript{10,39} $s^{a}(t - \theta^{a}_{l})$ is the systemic...
fluctuations, typically, the Mayer’s wave (resonance frequency of \( f_1 \)) and the breathing-related hemodynamic fluctuations (resonance frequency of \( f_2 \)), with a randomized delay of \( \theta_i \) for each channel and \( f_s \) is the sampling rate. Simulated signals for several channels and the correlation between them can be seen in Fig. 2(a). Figure 2(b) shows the original BHRF and the mixture signal from Eq. (5).

In all the simulations, \( w_1^f = 1 \) while \( w_1^j = w_2^f = w_3^f \) and \( w_1^j = \{0.1, 0.2, \ldots , 1\} \).

**2.1.2 Functional connectivity analysis**

We have decided to use a rectangular probe geometry to simulate the fNIRS signals, as shown in Fig. 3.

We decided to compute the FC using a PC-based analysis. PC provides a relationship between two variables after removing the overlap from both variables. The diagram in Fig. 4 depicts PC between time series 1 and 2 in the presence of a third time series 3. The PC coefficient between 1 and 2 after removing the influence of 3 \((r_{1,2|3})\) is as follows:

\[
r_{1,2|3} = \frac{r_{1,2} - r_{1,3}r_{2,3}}{\sqrt{(1 - r_{1,3}^2)}\sqrt{(1 - r_{2,3}^2)}}. \tag{8}
\]

This equation can be generalized to compute PC between any two channels \((i, j)\) as \(r_{i,j|k}\) in the presence of a common influencer \((k)\).

Since the fNIRS signal model assumes a linear addition of the BHRF with systemic fluctuations and noise, we assumed a certain frequency band for the brain-related signals and systemic fluctuations.

**2.2 fNIRS Data**

The data were collected in an earlier study.\(^{24,40}\) Twelve healthy controlled subjects performed the computerized version of the color-word matching Stroop task.\(^{41}\) The subjects were asked to respond to 90 stimuli presented on a screen every 4 s, in groups of six stimuli within each block. Fifteen blocks were divided into five NS, five CS, and five IC type of stimuli and presented in a random fashion. There were 20 s of rest within each block. The subjects were asked to respond to stimuli by pressing either the left or the right button of the mouse based on a match or unmatch condition. fNIRS data were collected with a 16 channel dual wavelength continuous wave system with a sampling rate
of 1.7 Hz (ARGES Cerebro, Hemosoft Inc., Ankara, Turkey). The source–detector configuration is shown in Fig. 3 with a separation of 2.5 cm. Absorption data collected at each detector are converted to [Hb] and [HbO] via the modified Beer–Lambert’s Law. The protocol was approved by the Ethical Review Committee of Bogazici University.

### 2.2.1 Functional connectivity analysis

Both the [Hb] and [HbO] data from each channel are passed through a high pass filter (Butterworth, eighth order, \( f_c = 0.09 \) Hz) to obtain the HBO\(_k\) and HB\(_k\). The regressor used in PC-based FC analysis is obtained by averaging this signal over all the channels. Hence, \( \overline{\text{HBO}_k} = \sum \text{HBO}_k \) \( \overline{\text{HB}_k} = \sum \text{HB}_k \) is used to regress out the systemic physiological effects from the correlation of the unprocessed [HbO] ([Hb]) signals from two channels. Once the regressor is computed, NS, CS, and IcS parts are consolidated to form individual time series for these stimuli. The FC matrices computed for individual time series are thus termed as \( \mathcal{F}C_N, \mathcal{F}C_C, \) and \( \mathcal{F}C_I \).

### 2.2.2 Global efficiency

One of the aims of cognitive neuroscience is to investigate the neural correlates of cognition.\(^{42-44}\) Graph-based network analysis is the state-of-the-art methodology in brain connectivity. We considered the channels as a set of vertices \( V \) and the PC coefficients as assigned weights on the set of edges \( E \), between vertices to construct an undirected complete weighted graph \( G = (V, E) \).\(^{45-47}\) We investigated the FC graphs of the PCs of each channel for each stimulus type.

Global efficiency (GE) can be evaluated for a wide range of networks, including weighted graphs.\(^{46}\) Maximal possible GE occurs when all edges are present in the network. The GE value was computed by using the formulation of Latora and Marchiori,\(^{48}\) since it applies to work with weighted connectivity graphs. In this case, the GE is

\[
\text{GE} = \frac{1}{N(N-1)} \sum_{i \neq j \in E} \frac{1}{d_{ij}}.
\]

where \( d_{ij} \) is defined as the smallest sum of the physical distances throughout all the possible paths in the graph from \( i \) to \( j \).\(^{48} \) For weighted graphs, stronger connection weights correspond to shorter lengths. Equation (9) generates values of GE in the range of \([0,1]\).

### 3 Results

#### 3.1 Simulation Analysis

The simulated signals for various weights as in Eq. (5) were used to compute the FC matrices. The plots in Fig. 5(a) show a sample of such signals while the errors in estimating the FC matrices are given in Fig. 5(b).

The top plot in Fig. 5(a) depicts how difficult it is to observe the presence of the hemodynamic response from the raw [HbO] signal. Figure 5(b) shows how well the PC-based computation of the correlations is closer to real correlation values. As the weight of the systemic fluctuations increases and starts to dominate the whole signal, the accuracy of extracting the real correlation value decreases.

![Fig. 5(a)](attachment:Simulated HRF signals for different channels with a certain delay. A mixed signal simulation for channel 1 based on Eq. (5) with \( \phi(t) \) in red and the mixed signal \( h(t) \) in blue. The parameters of the simulation are \( w_{s} = 0.3, r = 1, f_{1} = 0.1 \) Hz, \( f_{2} = 0.25 \) Hz, \( f_{s} = 2 \) Hz, and \( s_{f}(t) = N(0,0.03) \). The first plot in (a) shows the raw HbO signal, second row Butterworth low pass filtered signals and last plot shows the averaged regressor signal HbOR after being passed through a Butterworth type high pass filter, (b) errors in estimating the FC matrices. \( E_{\text{CO}} \) is the error with respect to Pearson’s correlation coefficient, \( E_{\text{BW}} \) is the error after low pass filtering with a Butterworth filter, and \( E_{\text{PC}} \) is the error with respect to PC analysis. The numbers above the bars indicate the percent improvement in the accuracy (decrease in the errors) between the low pass filtered and PC-based connectivity matrices.

#### 3.2 Real Data Analysis

A sample of fNIRS data from subject 1 and the corresponding regressor signal (HBO\(_k\)) can be seen in Figs. 6(a) and 6(b). The high pass filter setting was set at 0.09 Hz.\(^{42-45}\) Once the regressor is obtained, the data are segmented into NS, CS, and IcS parts. The choice of these cutoff frequencies was based on the fact that Mayer’s wave is centered around 0.1 Hz with a slight variation from 0.09 to 0.11 Hz.\(^{52-55}\) So the choice for the cutoff frequency for the high pass filter was based on the lower end of the Mayer’s wave band. Note that no further filtering was applied to the raw data when computing the PCs. Once
the regressor is formed from the raw data, then this regressor was used in computing the PC pairs of channels of unfiltered fNIRS data. This way any smearing effect of the spikey artifacts due to low pass filtering was minimized.

### 3.2.1 Functional connectivity of real data

A comparison was made between the conventional FC method and the PC-based FC method. Conventional FC matrices were formed by computing the Pearson’s correlation coefficient of pairs of channels after each fNIRS channel was low pass filtered (in this case with a fourth order Butterworth low pass filter with a cutoff frequency of 0.08 Hz).<sup>49–51</sup> Figures 6(c) and 6(d) show the average of FC matrices for IcS calculated with Pearson’s correlation to low pass filtered data (n = 12 subjects, \( C_{ij} \)) and (d) \( C_{ij} \) matrix calculated by PC.

Note that there are high correlations (over 0.8) in almost all the channel pairs for the Pearson’s coefficients [Fig. 6(c)] but not that diffusely scattered for the PC matrix [Fig. 6(d)]. This can be attributed to the fact that there are underlying systemic fluctuations embedded in each channel that dominate the overall correlation between two channels as hypothesized by the signal model in Sec. 2.1.1. This dominance seems to be cleared away in the FC matrix computed via the PC method, as shown in Fig. 6(c).

Table 1 compares numerically the FC matrices for three different stimuli computed by the Pearson’s correlation, PC, and thresholded PC approach in which the common regressor \( \text{HBO}_2^S \) (S is NS, CS, or IcS) is used.

These results elucidate the fact that there is a strong dominance of the background activity present in the signals that represents itself as high correlations among channels. This strong dominance of the underlying background activity is also evident by the fact that as the stimulus gets more demanding, the average strength of correlation increases in Pearson’s coefficients but not in PC coefficients. This increase can be explained by a stimulation of the sympathetic system during activities demanding higher cognitive engagement that leads to an increase of heart rate, blood pressure, blood flow to the brain, and breathing rate. On the contrary, no significant increase

![Sample Data from Ch=2](image1)

![Regressor Signal: Average over all Detectors](image2)

![Average FC by Pearson Corr, n=12](image3)

![Average FC for IcS, No TH](image4)
is observed for the values obtained by the PC analysis (see second column of Table 1). The elimination of this dominance is also evident in the decrease of the average strength of the correlations, as shown in the last row of Table 1 (from 0.5285 to 0.4031, a 23% decrease).

### 3.2.2 GE values for real data

The FC matrices are usually thresholded at a certain cutoff level to leave only a percent of the strongest connections before the computation of GE values. A scan of threshold values leaving the strongest 15 to 25 values yielded the highest significance among three different stimuli at GE values for [HbO] FC matrices \((TH = 21)\) and \(TH = 24\) for [Hb]. Average of the thresholded [HbO] FC matrices computed after thresholding at these values is shown in Figs. 7(a)–7(d), while their average correlation values are shown in the last column of Table 1. The average of the correlation values in the last row of Table 1. The GE values computed from the FC matrices generated by regular correlation (Pearson’s correlation coefficient) of low pass filtered data (via Butterworth filter) did not show any significant differences for various types of stimuli \((p > 0.05)\). When the correlations are computed via the PC analysis, we observed an increase in the GE values as the cognitive task became more demanding. GE values apparently change with respect to the threshold used for the FC matrices. We swept the threshold values (TH) and observed the significance among the GE\(_{\text{NS}}\), GE\(_{\text{CS}}\), and GE\(_{\text{IC}}\) both for [Hb] and [HbO] values. The TH value shows that the highest significance was observed for \(TH = 21\) for [HbO], which corresponds to 8.75\% of the highest correlations when the diagonals are omitted \((21/240)\) and \(TH = 24\) for [Hb] data. In a study by Zhang et al.,\(^{10}\) a strong lateralization effect was observed, favoring the flow of information to the right side. We grouped the detectors into four areas, where L (left) corresponds to the GE computed from detectors from the FC matrices 1 through 8 \([i = 1 \ldots 8, j = 1 \ldots 8\) in Eq. (9)], R (right) for detectors from 9 through 16 \([i = 9 \ldots 16, j = 9 \ldots 16\) in Eq. (9)], IH corresponds to interhemispheric connectivity and the GE was computed from FC matrices of 1 through 8 with 9 through 16th detectors \([i = 1 \ldots 8, j = 9 \ldots 16\) in Eq. (9)], as shown with dark squares in Fig. 7(d). Whole (W) corresponds to the GE computed from the full FC

![Fig. 7](image-url)

**Fig. 7** Average of FC matrices over all subjects for (a) NS, (b) CS, (c) IcS, and (d) interference for IcS-NS are shown. Note that the interference values are in the range of -0.2 to 0.2.
Table 2. GE values (rounded) with respect to PFC areas for the three stimuli types. Standard deviations are discarded but statistical significance (p value) was computed by paired two-tailed t-test. n is the number of subjects included in this analysis.

<table>
<thead>
<tr>
<th>ST</th>
<th>L⁺</th>
<th>R</th>
<th>IH</th>
<th>W⁻</th>
<th>L⁺</th>
<th>R</th>
<th>IH</th>
<th>W⁻</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>0.57</td>
<td>0.51</td>
<td>0.43</td>
<td>0.12</td>
<td>0.51</td>
<td>0.48</td>
<td>0.34</td>
<td>0.10</td>
</tr>
<tr>
<td>C</td>
<td>0.53</td>
<td>0.52</td>
<td>0.42</td>
<td>0.12</td>
<td>0.49</td>
<td>0.46</td>
<td>0.36</td>
<td>0.11</td>
</tr>
<tr>
<td>IC</td>
<td>0.52</td>
<td>0.55</td>
<td>0.47</td>
<td>0.14</td>
<td>0.46</td>
<td>0.56</td>
<td>0.36</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Note: ST, stimulus type. * indicates p < 0.1, while † indicates p < 0.05.

Table 2 shows the change of GE with respect to different stimuli for different areas of the PFC. As the GE in the left both for [Hb] and [HbO] decreases with varying stimuli difficulty in a statistically nonsignificant manner, the GE in the right both for [Hb] and [HbO] shows a statistically significant increase. This finding is consistent with the literature where the Stroop interference affect was found to be localized bilaterally.30,41,42,56,57

Zhang et al.30 claim to see an increase in the information flow from left to right dlPFC during interference. Figure 7(d) and consecutively Table 2 lead us to speculate that connectivity strength increases in the left and right dlPFC, but the connectivity pattern is diffused in the left while being more focused for the right.

It should be mentioned that the number of subjects included in the analysis in Table 2 had to be trimmed to achieve statistical significance. Only the highest and lowest GE values obtained from [HbO] data are eliminated while a higher number of subjects had to be eliminated from [Hb] data to achieve same significance levels. This result is in line with the literature, where [HbO] data are favored over [Hb] in cognitive studies due to its higher sensitivity to cognition-related hemodynamic changes.

The average interference values computed from GE for [Hb] and [HbO] data are shown in Fig. 8(a). Interference of the behavioral data (as reaction rates) shows a positive (negative) correlation for the interference values of the [HbO] ([Hb]) data, as shown in Fig. 8(b).

4 Discussion

Systemic physiological fluctuations have a deteriorating effect on the accuracy of computing the correlation of fNIRS signals.28,58-61 Several studies have attempted to address this issue by either advanced signal processing techniques or improvements in instrumentation and probe design albeit at increased cost and complexity of engineering.2,62

The current study proposes the use of PC-based connectivity computation under the assumption that a far detector signal is contaminated with systemic fluctuations that cannot be separated or regressed out with advanced signal processing techniques due to overlaps in time and frequency domains. Hence, the only solution becomes a statistical means of computing the correlation between a pair of channels.

4.1 On the Accuracy of Low Pass Filtering

We have shown in the first two scenarios that filtering with a low pass filter does not improve the correlation estimation since we do not have access to the frequency characteristics of the systemic fluctuations. The shortcoming of a correlation computation after signals is low pass filtered, which can be proven by deductive reasoning as follows.

Assume that the i-th detector signal \( x_i(t) \) is modeled as the sum of the brain activity, \( s_i(t) \), and nuisance term, \( s^o(t) \), \( x_i(t) = s_i(t) + s^o(t) \) and j-th detector signal as \( x_j(t) = s_j(t) + s^o(t) \). Assuming that there is an overlap in the spectrum

\[
|s_i^p(f)| + |s^o(f)| = |x_i(f)|
\]

Fig. 8 (a) Average interference of GE \((GE_{inc} - GE_{in})\) from [Hb] \((n = 6)\) and [HbO] \((n = 10)\) data for the different areas of PFC (+ for p < 0.1, * for p < 0.05). (b) Correlation with the interference of the behavioral data (reaction rates \( RT_{inc} - RT_{in} \)) with interference of GE \((GE_{inc} - GE_{in})\) for the whole PFC.

Fig. 9 Spectral composition of the signal from the j-th detector, \( X_j(f) \), can be given as the sum of the spectrum of the band-limited brain signal, \(|S_i^p(f)|\), with the spectrum of the wide-band nuisance signal, \(|S^o(f)|\), which is the sum of systemic fluctuations and other types of noises. Typical filter response is superimposed on the last graph.
of these signals, as depicted in Fig. 9, then a low pass filter
applied to the sum signal (spectrum with $|X_i(f)|$) will inevitably
include a piece of the nuisance spectrum.

Since low pass filters have nonideal characteristics, a leakage
of unwanted signals will be present in the filtered signal. This
will result in inaccurate estimation of the correlation coefficient.
This same problem of a background activity dominating the
channels exists in EEG analysis, where due to volume conduc-
tion effect, channels are superimposed with this background
activity. This must be removed if accurate correlations are to
be computed. Our results show a deviation from the original
connectivity values even after signals are filtered with various
filter types. Error in FC matrices for the Butterworth type of
filter produces very similar yet lower error values compared
to no preprocessed data, as expected. As the contamination
weight increases (i.e., increase in $w_f$) so do the errors in com-
puting the correlation values, as shown in Fig. 5(b).

4.2 On the Accuracy of PC-based Approach

Since there is some sort of dependence between the brain signals
and the contaminating signals both in the time and frequency
domains, it would be extremely difficult to try to separate
them between each other with conventional signal processing
tools, such as digital filtering, PCA, or ICA. All these methods
require several conditions of independence in time or frequency
domain. Kalman filtering would work if there were access to
systemic fluctuation data with a detector placed closer to the
source (short detector), as has been proposed by several
investigators. Hence, the only solution if one wants to
compute the correlation under such dependence assumption
and a fixed source–detector separation is the use of a PC
approach, where one can compute the correlation under a con-
trolled variable. This is exactly what the simulation shows. Its
accuracy exceeds by far both the filtered scenarios by almost
two to three folds, as shown in Fig. 5(b). Yet even this approach
cannot reach a zero level accuracy due to its use of a filter (this
time a high pass filter used when computing the nuisance signal
that will be used as the controlling variable). Hence, the success
of this method depends on how good the nuisance signal can be
provided to the algorithm. One solution is the use of a short
detector, which will pick up signals only from the skin.
Similarly, the use of a pulse oximetry signal even from the fin-
gertip might prove to be useful as long as one can account for the
delays in the systemic activity. Similar to the low pass filter’s
performance, the PC-based approach starts to fail as the con-
tamination level increases.

One alternative approach could be the use of this common
nuisance signal ($\text{HBO}_2$) in regressing it from the main detector
signal using a general linear model (GLM) based approach. This
might be possible but it will introduce other computational steps
before actually computing the correlation (first the GLM
approach, then reconstruct the HBO signal with the $\beta$’s
that are statistically inferred from the GLM approach, then compute
the correlation). Hence, it will undoubtedly introduce more errors
even due to the use of numerical methods.

4.3 Real Data Analysis

Having shown that if one has access to only far detector data and
no means to measure some sort of systemic activity, then PC-
based connectivity analysis is the best choice. Our results show
that with conventional means of computing, the connectivity
yielded statistically nonsignificant findings in our fNIRS con-
nectivity analysis. A very significant difference is observed
among the GE for the three different stimulus types, as
shown in Table 2. Several studies have shown a right dominance
for the Stroop activity. Our results also confirm a shift toward
right dIPFC connectivity as the stimulus becomes more chal-
lenging [from neutral stimulus (NS) to incongruent stimulus
(IC)], as shown in Table 2. As the GE in the left PFC decreases
for increasing task complexity, the right PFC picks up and
the interhemispheric GE increases, as shown in Fig. 8(a) although
not significantly. There is a clear dominance of the left dIPFC, as
indicated by many activation and connectivity studies.

Astolfi et al. showed a bilateral connectivity in the Stroop
task with a “predominance of outflow from right premotor
and prefrontal cortical areas.” In both the [Hb] and [HbO] GE
results, we see a drop in the incongruent stimuli compared to
the neutral stimuli in the left side and a consequent increase in
the interhemispheric and whole head interference results
[Fig. 8(a)]. This can be interpreted as the aggregation of shorter
paths to one longer path (leading to an increase in GE) for the
[HbO] signal.

Similar to some fMRI studies, there is a negative correlation
between the interference observed in reaction rates and the inter-
ference computed by GE values, as shown in Fig. 8(b), for
[HbO] at the left and interhemispheric connectivity metrics.
A significant positive correlation is observed for the [HbO]
for whole head connectivity analysis ($r = 0.729, p = 0.0259$)
while a significant negative correlation ($r = -0.745, p = 0.0213$) is observed for interhemispheric connectivity from
[Hb] data, as shown in Fig. 8(b). A positive correlation between
the interference computed for the reaction rates and GE values
means that as the cognitive challenge increases, sequential short
paths are replaced by one long path, a direct connection between
distant areas. An exact opposite is observed for the [Hb] inter-
ference, where the negative correlation means that a smaller
interference in reaction times gives higher interference in GE
values. An explanation can be as follows: as the cognitive chal-
lenge increases, brain regions working in coherence increase,
leading to a simultaneous demand of oxygenated blood to those
regions (hence an increase in the GE observed for [HbO]), yet
the venous side is not necessarily as coordinated leading to a
drop of efficiency. Neurovascular coupling literature has united
on the finding that both neurons and astrocytes may lead to a
vasodilation of the arteriole smooth muscle cells in response
to glutamate releases during neuronal activation. We
might also infer that [Hb] activity is mostly regulated on a
regional basis in a passive way and independent of the unifica-
tion required for responding to a cognitive challenge, much
similar to a balloon effect. That type of demand can be
explained only by the increase in [HbO] activation and its
GE metrics.

5 Conclusions

This study aimed at proposing a PC-based FC analysis for
fNIRS [HbO] and [Hb] signals collected during a color word
matching Stroop task. GE values calculated from the connectiv-
ity metrics of [HbO] signals reveal that as the cognitive chal-
lenge increases, so does the GE value. The use of PC-based
analysis was preferred since the optode signals suffer greatly
from systemic physiological interference. We proposed a
method to eliminate its biasing effect when a correlation is com-
buted between a pair of channels. This way the reliability of the
connectivity analysis is improved and the results are consequently more accurate.

Disclosures
There are no conflicts of interest to disclose.

Acknowledgments
This work is sponsored in part by TUBITAK Project Nos: 112E034 and 113E003. The data were collected as a part of projects funded by Bogazici University Research Fund Projects 04X102D, and 04S101.

References
56. I. Moser et al., “Right prefrontal brain activation due to stroop interference is altered in attention-deficit hyperactivity disorder—an fNIRS study,” Neuroimage 47, 569 (2009).

Ata Akin received his PhD in biomedical engineering from Drexel University in 1998. He holds his BS degree in electronics and telecommunications engineering from Istanbul Technical University. He serves as the dean of Faculty of Engineering at Acıbadem University, since 2015. His interests are in functional neuroimaging, fNIRS, mathematical physiology, and biodesign.