NIRS-EEG joint imaging during transcranial direct current stimulation: online parameter estimation with an autoregressive model
Mehak Sood, Pierre Besson, Makii Muthalib, Utkarsh Jindal, Stéphane Perrey, Anirban Dutta, Mitsuhiro Hayashibe

To cite this version:

HAL Id: lirmm-01402222
https://hal-lirmm.ccsd.cnrs.fr/lirmm-01402222
Submitted on 24 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
NIRS-EEG joint imaging during transcranial direct current stimulation: online parameter estimation with an autoregressive model

Mehak Sood* (Author)
Electronics and Communication Engineering
International Institute of Information Technology
Hyderabad, India

Pierre Besson* (Author)
EUROMOV
Université de Montpellier
Montpellier, France

Makii Muthalib (Author)
EUROMOV
Université de Montpellier
Montpellier, France

Utkarsh Jindal (Author)
Electronics and Communication Engineering
International Institute of Information Technology
Hyderabad, India

Stephane Perrey (Author)
EUROMOV
Université de Montpellier
Montpellier, France

Anirban Dutta* (Corresponding Author)
INRIA and Université de Montpellier
Montpellier, France

Present Address: Department of Biomedical Engineering
University at Buffalo SUNY
Bonner Hall #2151, Buffalo, USA 14228
Ph. No.: +1-716-645-2207
Email: adutta@ieee.org

Mitsuhiro Hayashibe* (Author)
INRIA and Université de Montpellier
Montpellier, France

*equal contribution, junior authors
*equal contribution, senior authors

Total text pages: 20
Total figures: 5
Total tables: 1
**Highlights**

- Transient coupling between EEG 0.5Hz-11.25Hz band-power and NIRS oxy-hemoglobin signal in the low frequency (≤0.1Hz) regime is proposed during tDCS.
- Time varying autoregressive (ARX) model is presented to track this transient coupling relationship during tDCS.
- Online parameter estimation of the ARX model is presented with a Kalman filter.
- Time varying poles associated with the ARX model were comparable across the healthy subjects during tDCS.
- Time varying zeros associated with the ARX model varied across the healthy subjects during tDCS.

**Abstract**— Background: Transcranial direct current stimulation (tDCS) has been shown to perturb both cortical neural activity and hemodynamics during (online) and after the stimulation, however mechanisms of these tDCS-induced online and after-effects are not known. Here, online resting-state spontaneous brain activation may be relevant to monitor tDCS neuromodulatory effects that can be measured using electroencephalography (EEG) in conjunction with near-infrared spectroscopy (NIRS). New Method: We present a Kalman Filter based online parameter estimation of an autoregressive (ARX) model to track the transient coupling relation between the changes in EEG power spectrum and NIRS signals during anodal tDCS (2mA, 10min) using a 4x1 ring high-definition montage. Results: Our online ARX parameter estimation technique using the cross-correlation between log (base-10) transformed EEG band-power (0.5-11.25 Hz) and NIRS oxy-hemoglobin signal in the low frequency (≤0.1Hz) range was shown in 5 healthy subjects to be sensitive to detect transient EEG-NIRS coupling changes in resting-state spontaneous brain activation during anodal tDCS. Conventional sliding window cross-correlation calculations suffer a fundamental problem in computing the phase relationship as the signal in the window is considered time-invariant and the choice of the window length and step size are subjective. Here, Kalman Filter based method allowed online ARX parameter estimation using time-varying signals that could capture transients in the coupling relationship between EEG and NIRS signals. Conclusion: Our new online ARX model based tracking method allows continuous assessment of the transient coupling between the electrophysiological (EEG) and the hemodynamic (NIRS) signals representing resting-state spontaneous brain activation during anodal tDCS.

**Keywords**—cross-correlation, tDCS, NIRS, EEG, ARX model, hemodynamics, neurovascular coupling, autoregulation
1. Introduction

Transcranial Direct Current Stimulation (tDCS) is a non-invasive way of modulating cortical excitability [1] and spontaneous brain activity [2]. It is a form of non-invasive brain stimulation (NIBS) which uses low direct electrical currents delivered directly to the brain area of interest through electrodes placed on the scalp. Anodal tDCS generally increases cortical neuronal excitability and activity [1], however the mechanisms are not known. Here, online resting-state spontaneous brain activation may be relevant to monitor tDCS neuromodulatory effects, specifically the relationship between cortical neural activity and regional hemodynamics, which can be respectively monitored using electroencephalography (EEG) in conjunction with functional near-infrared spectroscopy (NIRS).

During cortical activation, the electric currents from the excitable membranes superimpose at a given location in the extracellular medium and generate a potential which can be recorded as electroencephalogram (EEG) signals from scalp with excellent temporal resolution (milliseconds) and reasonable spatial resolution (usually around 6–9cm but can be improved to 2–3cm with Current Source Density estimates) [3]. Anodal tDCS after-effects have been shown to be limited to the alpha frequency (8–12 Hz) band of EEG signals [2]. The respective neural activity due to changes in cortical excitability requires a supply of oxygen and glucose that is delivered via the process of neurovascular coupling (NVC). NVC closely relates, spatially and temporally, to the regional hemodynamic response [4], which can be captured by functional magnetic resonance imaging (fMRI) [5][6] and functional NIRS neuroimaging [7][8]. Functional NIRS allows us to monitor the regional hemodynamic response non-invasively with reasonable spatial (~1cm, 0.5–2cm depth penetration [9]) and good temporal resolution [10]. Fast NIRS sampling at 250Hz is possible technically and relevant for fast and localized event-related optical signals [11]. However the time course of the regional hemodynamic response via NVC is much slower (in order of seconds) that limits the overall temporal resolution of the conventional NIRS method [12].

Preliminary studies have shown that anodal tDCS increases regional cerebral blood flow [13] and evokes neuronal and hemodynamic responses in the brain tissue [7], where widespread effects of tDCS on human NVC, vasomotor reactivity (VMR), and cerebral autoregulation have been found [14][15][16]. Some of these tDCS effects may be montage specific, such that a decrease in VMR after anodal tDCS [15] may be related to an extra cephalic montage (e.g., cathode placed on shoulder) affecting the brainstem [14]. Therefore, in the present study, we used a 4x1 ring high-definition tDCS (HD-tDCS) montage, where
HD-tDCS constrains the electric field between the active anode electrode at the center and the four surrounding return electrodes (~4cm apart) in a ring montage [17], to study local online resting-state spontaneous brain activation using the coupling between electrophysiological (EEG band power) and hemodynamic (NIRS oxy-hemoglobin: oxy-Hb) signals.

Prior EEG work has shown that spontaneous fluctuations in the alpha band (8–14 Hz) activity captures the momentary state of visual cortex excitability [18], which is the same frequency band that captured tDCS after-effects [2]. Furthermore, during awake resting-state joint EEG-NIRS measurements, slow (0.07-0.13 Hz) hemodynamic (oxy-Hb) oscillations have been shown to temporarily (duration ~100 sec) couple with slow EEG (alpha and/or beta power) power fluctuations in the sensorimotor cortex (SMC) areas [19]. The driving mechanisms responsible for these slow (0.07-0.13 Hz) hemodynamic oscillations is unclear; however, Nikulin et al [20] recently showed that monochromatic ultra-slow oscillations (MUSO) around 0.1 Hz in the EEG signals were related to the NIRS oxy-Hb signals, and the authors hypothesized that EEG MUSO represents an electric counterpart of the hemodynamic responses. Such fluctuations in the local cerebral oxygenation levels measured by oxygen polarography has been postulated to be driven by local cerebral autoregulation mechanisms [21], which plays an important role in oxygen delivery where impaired cerebral oxygenation oscillations may indicate a neurologic pathology [22]. Indeed, spontaneous dynamic cerebral autoregulation primarily reflected in the NIRS oxy-Hb (and total Hb) signal [22] can be a powerful tool to assess tDCS-induced cerebral vasodilation and autoregulation [23].

In principal accordance, we developed in our prior work [24] a phenomological model to capture the capacity of cerebral blood vessels to dilate during anodal tDCS due to neuronal (primarily synaptic) activity. Here, the coupling between the oxy-Hb time-series and log-10 (base-10) transformed (log_{10}) mean-power (in 0.5-11.25 Hz) time-course of EEG during anodal tDCS was shown [7], and the corticospinal excitability changes could be elucidated with NIRS-EEG joint-imaging [25]. Specifically, our prior work on four chronic (>6 months) ischemic stroke case series showed non-stationary effects of anodal tDCS on log_{10} mean-power of EEG within 0.5-11.25Hz frequency band that correlated with the NIRS oxy-Hb response [7]. An important aspect of the temporal profile of the NIRS oxy-Hb response was an initial dip at the beginning of anodal tDCS that corresponded with an increase in the log_{10} mean-power of EEG within 0.5-11.25Hz frequency band [7], which was explained by the ‘metabolic hypothesis’ of NVC [4]. However, an objective method for online tracking of the transient coupling relation between EEG and NIRS during tDCS could provide information on the momentary state of cortical excitability [18] and bidirectional interactions within the
neurovascular unit [26] that would be important for determining the dose of tDCS neuromodulation. Limitations of existing sliding window cross-correlation calculations [7][20][19] and conventional autoregressive-moving average models [27] that are sensitive to the subjective choice of the window length and step size for continuous monitoring of such non-stationary coupling relationship, have prevented development of such an online tracking model. Such window based cross-correlation calculations also have a fundamental problem in computing the phase relationship between NIRS and EEG, as the signal in the window is considered as time-invariant.

For the precise continuous online monitoring of the EEG and NIRS phase changes, it is required to use the framework of online tracking with an iterative update function which can assume the EEG and NIRS signals as time-varying. Therefore, in this methods paper, we developed and experimentally tested on 5 healthy subjects a computational autoregressive (ARX) model for online parameter estimation with a Kalman filter to track resting-state transient coupling relations between log10 mean-power EEG band (0.5-11.25 Hz) and NIRS oxy-Hb signals (≤0.1 Hz) acquired simultaneously from the left SMC during anodal HD-tDCS.

2. Materials and Methods

2.1 Subjects

Five healthy male subjects (age 36.6±13.13, all right-handed) volunteered to participate in this study after informed consent, and all experiments were conducted in accordance with the Declaration of Helsinki. The subjects had no known neurological or psychiatric history, nor any contraindications to tDCS. During the experiment, the subjects were comfortably seated with eyes-open in an armchair with adjustable height.

2.2 Experimental setup

The set-up of the HD-tDCS electrodes, EEG, and NIRS optodes was mounted on the surface of the scalp according to the 10/10 system (see Figure 1a) [28]. The anodal HD-tDCS (Startim®, Neuroelectrics NE, Barcelona, Spain) was configured in a 4x1 ring montage with the anode placed in the center (C3) in a region overlying the left SMC, and the return electrodes were placed approximately 4 cm away at FC1, FC5, CP5 and CP1 in the 10/10 system (see Figure 1b) [28]. In the rest of the document, ‘HD-tDCS’ means ‘anodal HD-tDCS’ since we only performed anodal tDCS for this human study.
NIRS: Measurements of changes in eyes-open resting-state oxy-Hb and deoxy-hemoglobin (deoxy-Hb) concentrations from the bilateral SMC were made from a 16-channel continuous-wave NIRS system (Oxymon MkIII, Artinis, Zetten, The Netherlands) at a sampling frequency of 10 Hz. The receiver-transmitter distance of 3cm was chosen based on Monte Carlo prediction of near-infrared light propagation [29] where sensitivity (S) at depth (d) is given by a “rule-of-thumb” formula, \[ S = 0.75 \times 0.85^d \] [30]. The receivers (Rx) were placed on the FC3 and CP3 for the left hemisphere, and FC4 and CP4 for the right hemisphere (see Figure 1a). Transmitters (Tx) were placed diagonally, i.e., at P1, P5, C1, C5, F5 and F1 for the left hemisphere and at P6, P2, C6, C2, F2 and F6 for the right hemisphere, as shown in Figure 1a.

EEG: Twenty-three channels of eyes-open resting-state raw unreferenced (active electrode) EEG signals were recorded (ActiveTwo, Biosemi B.V., The Netherlands). EEG artifacts related to tDCS are possible, e.g., due to slow time varying electrode impedance changes primarily between 0–2Hz [31], and requires adequate artefact rejection during signal conditioning. The twenty three EEG channels included AF3, AFz, AF4, Fz, FCz, FC2, FC6, FCC5h, FCC3h, Cz, C4, CCP5h, CCP3h, CPz, CP2, CP6, Pz, POz, O1, Oz, O2, Fp1, Fp2 in the 10/10 system (see Figure 1a) [28].

2.3 Experimental protocol

The experiment was divided into three epochs (pre-, during-, and after HD-tDCS) for the purposes of this computational work. Different epochs were marked by sending markers using TCP/IP to the Startim® (Neuroelectrics NE, Barcelona, Spain), and simultaneous acquisition of all the sensor (EEG and NIRS) signals were TTL-synchronized at the ActiveTwo (Biosemi B.V., Netherlands) and Oxymon MkIII (Artinis, Zetten, The Netherlands) systems using custom-written code in Matlab (Mathworks Inc., USA). The Pre-tDCS epoch had 3 minutes of rest to have a baseline acquisition. During the HD-tDCS epoch, 2mA current was delivered using PISTIM electrodes (Neuroelectrics NE, Barcelona, Spain) with 3.14cm² (1 cm radius) contact surface. HD-tDCS was delivered for 10 minutes with ramp-up of 30 seconds. All the synchronized sensor data files were saved and analyzed offline after the completion of the two sessions.

2.4 EEG and NIRS data pre-processing

The raw unreferenced (active electrodes) EEG data was preprocessed offline using EEGLAB open-source Matlab (Mathworks Inc., USA) toolbox [32]. The data was first
referenced to the ‘average reference’ and then artefactual epochs ("non-stereotyped" or "paroxysmal" noise such as those induced by generalized head and electrode movements, tDCS voltage ramping epochs, high amplitude random fluctuations in broadband EEG activity near HD-tDCS electrodes) were removed following visual inspection of the data. Then, further artefact rejection was performed with Independent Component Analysis (ICA) which provides a powerful tool for eliminating several non-brain artifacts from EEG data including low frequency tDCS-related artifacts \cite{32,31}. We followed the method outlined by Delorme et al. \cite{33} in EEGLAB: (1) detection thresholds were set such that roughly 10% of data trials were detected, (2) visually inspected the data trials marked as artefactual, and (3) then manually spectral thresholded outside of 5 standard deviations above the mean. Then, log10 transformed mean power (‘bandpower’ in Matlab) of ICA-cleaned EEG within 0.5-11.25 Hz frequency range was computed \cite{7}. The log (base-10) transformed (log10) EEG mean-power time-series was down sampled to 10 Hz to match the sampling frequency of the NIRS oxy-Hb data. Then, the baseline corrected NIRS oxy-Hb time-series as well as the log10 EEG mean-power time-series were 5th order Butterworth zero-phase digital low-pass filtered at 0.1 Hz cut-off frequency to isolate the slow oscillations \cite{19}.

### 2.5 Autoregressive (ARX) modeling

In this method study, to evaluate our Kalman Filter based method that allowed online parameter estimation of an autoregressive (ARX) model, we investigated NIRS oxy-Hb signal and log10 EEG band power within 0.5-11.25 Hz frequency range at the HD-tDCS site (anode at C3 and cathodes at FC1, FC5, CP5, CP1, see Figure 1a) from co-located EEG (FCC5h, FCC3h, CCP5h, CCP3h) and NIRS (T1-R2, T1-R1, T2-R2, T2-R1) sensors (see Figure 1a and 1b). The ARX model is a common method to represent output signals from an unknown system by using a linear combination of past output signal values and past input values. The respective sensor signals at the HD-tDCS site were averaged before being used for ARX modeling. We applied the ARX model approach to evaluate the degree of coupling between log10 EEG band (0.5-11.25 Hz) power and NIRS oxy-Hb dynamics in the slow oscillation regime (≤0.1 Hz) \cite{19}. Here, log10 EEG band (0.5-11.25 Hz) power was used as the input signal and NIRS oxy-Hb dynamics was used as an output signal under the hypothesis that tDCS-modulated neuronal activation induced tonic vasodilation can affect the cerebral oscillations in the slow frequency range (≤0.1 Hz) in healthy subjects \cite{23}. Our prior work developed a phenomological model to show the capacity of blood vessels to dilate during anodal tDCS due to neuronal (primarily synaptic) activity \cite{24}. 
2.5.1 Model Structure

The linear time variant system can be described by an autoregressive model with exogenous input (ARX) [34].

\[ A(z)y(t) = B(z)u(t) + e(t) \]  

(1)

with transfer function \( G(z) = \frac{B(z)}{A(z)} \), where

\[ A(z) = 1 + a_1z^{-1} + a_2z^{-2} + \ldots + a_lz^{-l} \]

\[ B(z) = b_1z^{-m} + b_2z^{-(l+m)} + \ldots + b_mz^{-(l+nm-1)} \]  

(2)

In equation (1), \( y(t) \) is the output and \( u(t) \) is the input at any discrete-time instant, \( t \), where \( t \) ranges over the integers. The \( z^{-1} \) is backward shift operator and \( z^{-1}y(t) \) is equivalent to \( y(t-1) \). Also, \( e(t) \) is the zero mean Gaussian white noise affecting the system. The model has \( (l + m + 1) \) parameters/coefficients in total, i.e., \((a_1, a_2, b_1, \ldots, b_m, n)\) where \( l \) is the number of poles, \( m \) is the number of zeros plus 1, and \( n \) is the number of input samples that occur before the input affects the output, also called the dead time in the system.

Substituting equation (2) in (1), and expanding, the output of the ARX model can be parameterized as

\[ y(t, \theta) = \sum_{i=1}^{l} a_iy(t-i) + \sum_{j=1}^{m} b_ju(t-j-n+1) + e(t) \]

(3)

where \( \theta = [a_1, a_2, b_1, \ldots, b_m]^T \) and the size of \( \theta \) depends on the complexity of the model. Thus, the selection of model order and structure becomes a crucial step in the estimation of unknown parameters in \( \theta \). The system identification toolbox in Matlab (The MathWorks Inc., USA) was used to find an optimal model order and structure \((l, m, n)\). The pre-tDCS recording with EEG power as input and NIRS oxy-Hb signal as output provided the estimate of the model order and structure \((l, m, n)\). A range of ARX polynomial models were investigated for all the subjects and the “simpler” model in terms of model order that gave comparable unexplained output variance (in %) to the “best” model for all the subjects was selected. Here, the elements of \( \theta \) are time varying as it relates the variations in EEG power to NIRS response. At a given time instant, \( t \), the model estimates are predicted using equation (3), assuming that the system is stationary or slowly time varying (e.g., during awake resting-state [19]) in the prediction horizon.

2.5.2 State Space Representation of ARX model
State space representation of the ARX model is required for the Kalman filter implementation \cite{34} where we considered an ARX $(l,m)$ model without the dead time since dead time was assumed constant and not estimated. The state space form was:

Process Equation:

\[ x_k = A x_{k-1} + B u_{k-1} \]  

(4)

Measurement Equation:

\[ y_k = C x_k \]  

(5)

where $k$ represents the current time step. In equation (4), the state vector at the current time step, $x_k = [x_1 \ldots x_q]^T$, $q = \max\{l,m\}$, and $u_{k-1}$ is the model input at the previous time step. $A \in \mathbb{R}^{(q \times q)}$ relates the state vector at previous time step, $x_{k-1}$, to that of current time step, $x_k$. $B \in \mathbb{R}^{(q \times l)}$ relates the model input at the previous time step, $u_{k-1}$, to the state vector at the current time step, $x_k$. The measurement of the system output at the current time step is given by $y_k$. $C \in \mathbb{R}^{(l \times q)}$ relates the state vector at the current time step, $x_k$, to the measurement at the current time step, $y_k$. Here, $A,B,C$ may change with each time step of iteration, but in this study we assumed them constant for simplification.

2.5.3 Kalman filter for ARX online parameter estimation

A Kalman filter is an efficient filter which consists of a set of mathematical equations that implement a predictor-corrector type estimator. A Kalman filter is optimal in the sense that it minimizes the estimated error covariance—when some conditions are met. It is recursive in nature and can efficiently estimate the internal states and parameters of a discrete time system from a series of noisy measurements. Here, the coefficients relating the past measurements of NIRS oxy-Hb signal as output and past measurements of EEG power as input are the model parameters that are to be identified. Therefore, the state vector, $x$, was augmented with the unknown model parameters, $\theta$, that are to be identified. Here, by regarding the unknown model parameters as elements of the state vector, the basic Kalman filter algorithm was followed for the estimation of the state vector as well as the model parameters. Consequently, the meta-state vector, $w_k = [x_k; \theta_k]^T$, was created for the basic Kalman filter algorithm where $k$ indicates the current time step. The model parameters, $\theta$,
were assumed to be locally time-invariant when compared to the process. Consequently, the augmented Kalman filter system was given by,

Process Equation:
\[ w_k = F(w_{k-1}, u_{k-1}) \]  
(6)

where,
\[ F(w_{k-1}, u_{k-1}) = \begin{bmatrix} f(x_{k-1}, u_{k-1}) \\ \theta_{k-1} \end{bmatrix} \]

and \( f \) is a function of the state vector, \( x_{k-1} \), and the model input, \( u_{k-1} \), at the previous time step.

Measurement Equation:
\[ y_k = Hw_k \]  
(7)

where,
\[ H = \begin{bmatrix} C & 0^{(1+m)} \end{bmatrix} \]

Now, the recursive estimation of the state space model can be performed using the prediction and correction steps of the basic Kalman filter algorithm where,

Prediction Step:
\[ \hat{w}_k^- = F(\hat{w}_{k-1}^-, u_{k-1}) \]  
(8)
\[ P_k^- = D_k P_{k-1} D_k^T + Q_{k-1} \]  
(9)

Correction Step:
\[ K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \]  
(10)
\[ \hat{w}_k = \hat{w}_k^- + K_k (y_k - H_k \hat{w}_k^-) \]  
(11)
\[ P_k = (I - K_k H_k) P_k^- \]  
(12)

Here, in the equation (8) of the prediction step, the a priori estimate of the meta-state vector at the current time step, \( \hat{w}_k^- \), is given by the process equation (6) using the a posteriori estimate of the meta-state vector at the previous time step, \( \hat{w}_{k-1}^- \). Also, the a posteriori
estimate of the error covariance from the previous step, $P_{k-1}$, is propagated as per equation (9) to the \textit{a priori} estimate of the current step, $P_k^{-}$, where $Q_{k-1}$ is a diagonal matrix containing the process noise covariance, and $D_k$ is the process Jacobian with respect to the meta-state vector elements,

$$D_{[i,j]} = \frac{\partial F_{[i,j]}}{\partial \hat{w}_{k-1}, u_{k-1}}$$  \hspace{1cm} (13)$$

In the equation (10) of the correction step, $K_k$ is the Kalman filter gain which is computed from the \textit{a priori} estimate of the error covariance of the current step, $P_k^{-}$, the scalar measurement noise covariance, $R_k$, and the mapping matrix, $H_k$, from the equation (7) computed at the current time step, $H_k$. Using $K_k$, the Kalman filter gain, the \textit{a posteriori} estimate of the meta-state vector at the current time step, $\hat{w}_k$, as well as the \textit{a posteriori} estimate of the error covariance at the current time step, $P_k$, are computed in equations (11) and (12) respectively. However, the limitation in this approach is the degradation of Kalman filter’s performance in estimating time varying parameters as the basic Kalman filter algorithm refers to the entire history of past measurements [34]. Since tDCS may lead to transients in NIRS oxy-Hb signal, therefore tracking of this becomes important with the model parameters. In order to track this transient activity, a forgetting factor lambda, $\lambda$, was introduced in the basic Kalman filter algorithm [34],

$$P_k^{-} = D_k P_{k-1} D_k^T / \lambda$$  \hspace{1cm} (14)$$

$$K_k = P_k^{-} H_k^T (H_k P_k^{-} H_k^T + \lambda)^{-1}$$  \hspace{1cm} (15)$$

When the forgetting factor is closer to 1 then the filter will forget fewer past measurements where a trade-off between the smoothness of tracking and the lag in detecting the changes in model parameters should be considered [34]. Usually, $\lambda \in [0.9, 1]$ is suitable for most applications.

### 2.5.4 Validation of Time-Varying ARX Model Estimation with a Kalman Filter

The elements of the meta-state vector, $W_k = [x_k; \theta_k]^T$, were initialized to zero. The initial output estimate was also set to zero. The estimate of the error covariance was initialized to identity matrix, $P_0 = I$. In simulation, the invariant parameter tracking was evaluated first with the Kalman filter to investigate the stability of the model. Then, in order
to investigate the Kalman filter’s robustness to the time varying phenomenon, the model parameters were slowly changed to simulate tracking of time-varying properties of non-stationary signals. The advantage of simulation is that the true parameters are known and therefore can be compared with the estimated ones. The ARX model structure was chosen as \( (l = 3, m = 3, n = 1) \) for simulations. Thus, six parameters, \( \theta = [a_1, a_2, a_3, b_1, b_2, b_3]^T \), were estimated via the Kalman filter algorithm in simulation. Here, a pseudo random binary sequence was chosen as the model input [34].

Following validation using simulated signals, the time-varying ARX model estimation with a Kalman filter was applied to the 5 healthy subjects EEG-NIRS data and the online tracking results were compared with that from conventional sliding window cross-correlation calculations [19]. Here, a moving window size of 100sec with an overlap of 0.1 second and the number of lag of ±20sec (Matlab function ‘crosscorr’) was selected for conventional sliding window cross-correlation calculations based on prior work [19], which identified short-lasting coupling of ~100sec duration with <20sec lead/lag time. Cross-correlation values were set to zero within ±3 standard deviations for the estimation error assuming the signals to be uncorrelated.

3. Results

Figure 2a shows that the simulated [29] electric field in the left SMC gray matter surface during anodal HD-tDCS. The electric field simulation was limited to the left SMC, as mentioned in a related prior work [35], and no affect of the electric field on the brainstem structures was found. Figure 2b shows the simulated [29] measurement sensitivity distribution of the NIRS optodes at the left SMC gray matter surface where the penetration depth was roughly 1.5cm below the scalp between a NIRS emitter and receiver in concurrence with the “thumb rule” (penetration depth is approximately one-half of the distance between NIRS emitter and receiver [36]).

During validation of our online parameter estimation with a simulated time-varying ARX model, the mean absolute error of the parameter estimates with respect to true parameter value were 0.008±0.0130, 0.007±0.0105, 0.004±0.0036, 0.003±0.0185, 0.003±0.0174, 0.004±0.0204 for the parameters \( a_1, a_2, a_3, b_1, b_2, b_3 \), respectively.

After validating our Kalman filter approach for time-varying ARX model estimation, the system identification toolbox in Matlab (The MathWorks Inc., USA) was used to find the best model order from the pre HD-tDCS experimental data. The ARX model
was found to be the simplest model across all the 5 subjects. Therefore, the ARX model \((l = 4, m = 5, n = 5)\) with 4 poles, 6 zeros, and a constant dead time, \(n = 5\) (i.e., total 0.5sec dead time for 5 sample time-periods of 0.1sec), was used for the online tracking of the coupling relationship between \(\log_{10}\) EEG band (0.5-11.25Hz) power as input and NIRS oxy-Hb signal as the output in the slow frequency range (\(\leq 0.1\) Hz). A small (0.5sec) constant dead time meant that we tracked primarily in-phase and out-of-phase coupling between NIRS oxy-Hb signal (as output) and \(\log_{10}\) EEG band (0.5-11.25 Hz) power signal (as input) which has been shown to be present during awake resting-state measurements [19].

In accordance, Figure 3a shows with an illustrative example (subject 5) the short-duration (~100sec) in-phase and out-of-phase coupling between NIRS oxy-Hb signal and \(\log_{10}\) EEG band (0.5-11.25 Hz) power signal around 294sec and 428sec. The ARX parameters values found for the 5 subjects at the start of HD-tDCS (pre HD-tDCS) and after the end of HD-tDCS session (post HD-tDCS) are given in Table 1. Also, the Prediction Error (Measured signal - Modeled signal) during ARX online tracking is shown in Figure 3b for the same illustrative example (subject 5) and the root mean square error (RMSE) between the measured and modeled NIRS oxy-Hb signal (output) during HD-tDCS is presented in Table 1 for all the subjects.

The variation in the ARX parameters due to this change in the cross-correlation between \(\log_{10}\) EEG band (0.5-11.25 Hz) power and NIRS oxy-Hb signal was compared with sliding cross-correlation calculations [19][15], as shown in Figure 4a, for the same subject. In Figure 4b, transients in the time-series of zeros of the ARX model, found around 294sec and 428sec (also see the zoomed views in Figure 3a), are shown with conventional cross-correlation function (and its estimation error bounds) that measured the similarity between NIRS oxy-Hb signal and 20sec shifted (lagged) copies of EEG band (0.5Hz-11.25Hz) power signal within the 100sec window centered at 294sec (see top panel of Figure 4b) and 428sec (see bottom panel of Figure 4b).

4. Discussion

This is the first study to develop an online autoregressive (ARX) model based tracking method to continuously monitor the transient coupling between the electrophysiological (EEG) and the hemodynamic (NIRS) signals during anodal HD-tDCS. In 5 healthy subjects, we presented a Kalman Filter based online parameter estimation of an ARX model with 4 poles, 6 zeros, and a constant dead time of 0.5sec, that was found suitable
(RMSE=0.11±0.06 shown in Table 1) to capture the transfer function from log\textsubscript{10} EEG band (0.5-11.25 Hz) power alterations to NIRS oxy-Hb signal changes in the slow frequency range (≤0.1 Hz) during anodal HD-tDCS of the left SMC. This online ARX parameter estimation using a Kalman filter provided a more convenient method to capture this time-varying EEG to NIRS transfer function when compared to sliding cross-correlation calculations [19], which suffers a fundamental problem in computing the phase relationship as the signal in the window is considered as time-invariant, and the choice of the window length and step size are subjective.

During EEG signal conditioning, removal of the tDCS related artifacts or noise is important for subsequent ARX parameter estimation. For this purpose, we endeavoured to identify tDCS-related drift artifacts (primarily between 0–2Hz [31]) in the EEG time series for removal using the ICA function in EEGLAB by spectral thresholding outside of 5 standard deviations above the mean. We used a Kalman Filter based online ARX parameter estimation that does not require selection of a window length and step size, and is robust to Gaussian measurement noise [37]. In fact, Kalman filter is the best linear estimator given the mean and the standard deviation of the noise. In case of non-Gaussian measurement noise with its covariance matrices found from phantom studies [31], the noise can also be appropriately modeled using a more advanced (unscented) Kalman Filter [37] to make the parameter estimation insensitive to the modeled noise. Other advanced algorithms have also been proposed to handle non-Gaussian noise, e.g., recursive algorithm [38] and Gaussian sum filter [39]. The ARX parameter estimation approach presented in this paper is amenable to such advanced Bayesian estimation techniques when necessary based on the data quality.

Table 1 shows the transfer function at the start and at the end of HD-tDCS where the poles \((a_1,a_2,a_3,a_4)\) are the roots of the denominator of the transfer function. Indeed, the poles associated with the output side, i.e. NIRS oxy-Hb signal, have a direct influence on the dynamic properties of the system which were comparable across the 5 healthy subjects pre- and post- HD-tDCS. The zeros \((b_1,b_2,b_3,b_4,b_5)\) are the roots of the numerator of the transfer function which are associated with the input side, i.e., log\textsubscript{10} EEG band (0.5-11.25 Hz) power. The zeros varied during HD-tDCS, as shown in Table 1 (also, Figure 4a for subject 5), which indicated HD-tDCS effects on the log\textsubscript{10} EEG band (0.5-11.25 Hz) power. In accordance, the standard deviation when compared to the mean, i.e. the the coefficient of variation in Table 1, is lower for poles \((a_1,a_2,a_3,a_4)\) and higher for zeros \((b_1,b_2,b_3,b_4,b_5)\) that shows the slow oscillation rhythm of the NIRS oxy-Hb signal was more consistent across the subjects than
the slow oscillation rhythm of the EEG band power signal during HD-tDCS. When we visually compared the pre- and during- HD-tDCS NIRS oxy-Hb signal and EEG band power signal for a single subject (subject 5) in Figure 4a, this observation is consistent where – (1) EEG signal rhythm was more affected by HD-tDCS, such as in-phase and out-of-phase alterations, when compared to a consistent NIRS oxy-Hb rhythm and (2) NIRS amplitude was more altered by HD-tDCS, e.g. at the beginning (0-200sec) of HD-tDCS, than its rhythm during HD-tDCS – which needs further investigation in the future for physiological interpretation.

These observation from Figure 4a that the zeros of the ARX model were substantially modulated at the beginning of the HD-tDCS that corresponded with the initial response of oxy-Hb and deoxy-Hb to HD-tDCS (shown in Figure 5) points to ‘metabolic hypothesis,’ i.e., HD-tDCS-evoked local neuronal activity increased energy metabolism which then drives vasodilation increasing the regional cerebral blood flow (CBF) and oxy-Hb drastically [4]. Indeed, an initial dip in oxy-Hb (and an increase in deoxy-Hb) due to oxygen metabolism can be seen in Figure 5 that was also found in our prior work on stroke subjects [7]. Not only at the beginning (0-200sec) of HD-tDCS, but also following (>200sec) that initial response during HD-tDCS, our online parameter estimation continuously identified transient coupling relations between EEG band power and NIRS oxy-Hb signals primarily with the alterations in the zeros \( b_1, b_2, b_3, b_4, b_5 \) of the transfer function. The cross-correlation (zero lag) between \( \log_{10} \) EEG band (0.5-11.25 Hz) power (as input) and NIRS oxy-Hb signal (as output) in the slow frequency regime (\( \leq 0.1 \) Hz) fluctuated between negative correlation (top panel of Figure 4b) and positive correlation (bottom panel of Figure 4b) in the healthy subjects (Subject 5 shown in Figure 4) that was also found in our prior work on stroke subjects [7]. Here, after that initial response (0-200sec) possibly related to ‘metabolic hypothesis,’ close to zero lag cross-correlation supports a ‘neurogenic hypothesis,’ i.e., regional CBF and energy metabolism are driven in parallel following that initial response to HD-tDCS (i.e., >200sec) when vasodilation/constriction are driven in a feedforward manner [10]. A cross-correlation function (see Figure 4b) also demonstrated statistically significant (±3 standard deviations for the estimation error assuming the signals to be uncorrelated) cross-correlation at other lags which may be related to the transient bidirectional interactions [26] between the neuronal and the hemodynamic components during HD-tDCS.

Bidirectional interactions between the neuronal and the hemodynamic components can be caused due to indirect (via neuro-glio-vascular unit - ‘metabolic hypothesis’) and direct
(e.g. feedforwad ‘neurogenic hypothesis’, glial effects [40]) effects of tDCS where glucose utilization in astrocytes can result in an increase in the extracellular lactate levels which when taken up by the neurons can promote membrane depolarization, excitability alterations, and its after-effects [41]. Also, direct tDCS effects on endothelial cell-dependent responses is possible [42] besides neuro-glio-vascular unit mediated responses [26]. Although low-dimensional physiological models of the neuro-glio-vascular unit may be necessary to elucidate such bidirectional interactions from multimodal neuroimaging data [43], nevertheless, system identification techniques presented in this paper possibly with a time varying (positive and negative) dead time are required for online tracking of this transient coupling between EEG band power and NIRS oxy-Hb time-series for titrating tDCSs’ online excitability alterations and after-effects [25] for its closed-loop control [26]. Moreover, subject-specific alterations of ARX poles and zeros with different dead time may be relevant for diagnosing neurovascular dysfunction in clinical populations since electrophysiological signatures of the resting brain state [44] may be dysfunctional in cerebrovascular occlusive disease [45], which needs to be investigated using computational models [46] in future works.

In conclusion, our novel Kalman Filter based method was shown for online parameter estimation of an ARX model between EEG band power and NIRS oxy-Hb signals from the left SMC during resting-state anodal HD-tDCS. In future studies, this same approach will be used to capture tDCS induced transient EEG-NIRS coupling during task related brain activation as well as other transient coupling relationships, e.g., to investigate dynamic cerebral autoregulation, cerebrovascular reactivity, and neurovascular coupling in health and disease [27].

**Author Contributions**

AD MM SP contributed to the conception of this scientific investigation while MH AD contributed to the conception of the computational method – AD and MH equal contribution. MS substantially contributed to the data analysis alongwith UJ while PB substantially contributed to the data acquisition alongwith MM – MS and PB equal contribution. AD, MM, MH, SP have contributed to the interpretation of the results, writing and reviewing of the manuscript. All authors have approved the final version prior to submission and are in agreement.
Acknowledgement

MS and UJ were supported by Franco-Indian INRIA-DST project funding. MM and this experimental work was funded by the LabEx NUMEV (ANR-10-LABX-20).

References


Figure 1. a) Set-up of HD-tDCS electrodes, EEG, and NIRS optodes in 10/10 system. b) NIRS-EEG montage at the HD-tDCS site (red ellipse) overlying the left primary sensorimotor cortex (SMC).

Figure 2. a) Electric field (V/m) estimated at the gray matter surface due to 2mA anodal HD-tDCS. b) Measurement sensitivity distribution of the NIRS optodes at the gray matter surface.
Figure 3. a) An illustrative example (subject 5) showing the log_{10} EEG band (0.5-11.25 Hz) power as the input (in black), NIRS oxy-Hb signal that was measured (in blue) as well as the predicted NIRS oxy-Hb signal using the ARX online tracking method (in dotted red). b) Prediction Error (Measured signal - Modeled signal) during ARX online tracking is shown. Inset: Short-duration (~100sec) in-phase and out-of-phase coupling between the NIRS oxy-Hb signal and EEG band (0.5-11.25 Hz) power signals around 294sec and 428sec are shown with zoomed view that corresponded with the transients in the ARX model parameters (see also Figure 4a).
Figure 4. a) An illustrative example (subject 5) of the correspondence between the ARX parameter tracking (time series of the poles and zeros shown in the top two panels) and the sliding cross-correlation function (bottom panel) for NIRS oxy-Hb signal as the output and log_{10} EEG band (0.5-11.25 Hz) power as the input in the slow frequency range (< 0.1 Hz). Transients (shown with ellipse) in the time-series of zeros (middle panel) can be found around 294sec and 428sec that corresponded with that found from conventional sliding cross-correlation function (bottom panel). b) Cross-correlation function shows a short-duration (window length=100sec) negative coupling around 294sec (top panel) while positive coupling occurs around 428sec (bottom panel). Cross-correlation function measured the similarity between NIRS oxy-Hb signal and up to 20sec shifted (lagged) copies of EEG band (0.5-11.25 Hz) power signal in the slow frequency regime (< 0.1 Hz) where the dashed horizontal lines show ±3 standard deviations for the estimation error assuming the signals to be uncorrelated.
Figure 5. An illustrative example (subject 5) showing NIRS oxy-Hb (in blue) and NIRS deoxy-Hb (in black) signals. Small increase in deoxy-Hb signal and decrease in oxy-Hb signal was found at the beginning due to anodal HD-tDCS-evoked local neuronal activity increasing oxygen metabolism, that then caused vasodilation increasing regional cerebral blood flow and an increase in oxy-Hb signal, which then reduced deoxy-Hb signal, i.e., ‘metabolic hypothesis’ of neurovascular coupling.
Table 1
(Std. Dev. is standard deviation, CV is coefficient of variation)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Subject 1</th>
<th>Subject 2</th>
<th>Subject 3</th>
<th>Subject 4</th>
<th>Subject 5</th>
<th>Mean±Std. Dev. (CV = Std. Dev/mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre HD-tDCS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_1$</td>
<td>3.24</td>
<td>3.25</td>
<td>3.24</td>
<td>3.23</td>
<td>3.21</td>
<td>3.23±0.02 (±0.006)</td>
</tr>
<tr>
<td>$a_2$</td>
<td>-3.73</td>
<td>-3.77</td>
<td>-3.76</td>
<td>-3.74</td>
<td>-3.72</td>
<td>-3.74±0.02 (±0.005)</td>
</tr>
<tr>
<td>$a_3$</td>
<td>1.78</td>
<td>1.77</td>
<td>1.79</td>
<td>1.72</td>
<td>1.71</td>
<td>1.75±0.04 (±0.023)</td>
</tr>
<tr>
<td>$a_4$</td>
<td>-0.25</td>
<td>-0.26</td>
<td>-0.25</td>
<td>-0.24</td>
<td>-0.29</td>
<td>-0.26±0.02 (±0.077)</td>
</tr>
<tr>
<td>$b_1$</td>
<td>0.025</td>
<td>0.018</td>
<td>0.096</td>
<td>0.001</td>
<td>0.270</td>
<td>0.08±0.11 (±1.375)</td>
</tr>
<tr>
<td>$b_2$</td>
<td>0.016</td>
<td>0.001</td>
<td>0.061</td>
<td>0.664</td>
<td>0.001</td>
<td>0.15±0.29 (±1.933)</td>
</tr>
<tr>
<td>$b_3$</td>
<td>0.001</td>
<td>0.182</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.04±0.08 (±2.000)</td>
</tr>
<tr>
<td>$b_4$</td>
<td>0.001</td>
<td>0.147</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.03±0.07 (±2.333)</td>
</tr>
<tr>
<td>$b_5$</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.01±0.00 (±0.000)</td>
</tr>
<tr>
<td>post HD-tDCS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_1$</td>
<td>3.26</td>
<td>3.19</td>
<td>3.25</td>
<td>2.99</td>
<td>2.76</td>
<td>3.09±0.21 (±0.068)</td>
</tr>
<tr>
<td>$a_2$</td>
<td>-3.79</td>
<td>-3.72</td>
<td>-3.74</td>
<td>-3.27</td>
<td>-2.98</td>
<td>-3.5±0.36 (±0.103)</td>
</tr>
<tr>
<td>$a_3$</td>
<td>1.79</td>
<td>1.84</td>
<td>1.73</td>
<td>1.59</td>
<td>1.62</td>
<td>1.71±0.10 (±0.059)</td>
</tr>
<tr>
<td>$a_4$</td>
<td>-0.26</td>
<td>-0.32</td>
<td>-0.24</td>
<td>-0.31</td>
<td>-0.40</td>
<td>-0.31±0.06 (±0.194)</td>
</tr>
<tr>
<td>$b_1$</td>
<td>-0.013</td>
<td>0.008</td>
<td>-0.014</td>
<td>-0.134</td>
<td>0.0752</td>
<td>-0.02±0.08 (±4.000)</td>
</tr>
<tr>
<td>$b_2$</td>
<td>0.029</td>
<td>-0.019</td>
<td>0.043</td>
<td>0.557</td>
<td>-0.006</td>
<td>0.12±0.25 (±2.083)</td>
</tr>
<tr>
<td>$b_3$</td>
<td>-0.003</td>
<td>0.007</td>
<td>-0.029</td>
<td>-0.743</td>
<td>-0.225</td>
<td>-0.20±0.32 (±1.600)</td>
</tr>
<tr>
<td>$b_4$</td>
<td>-0.028</td>
<td>0.009</td>
<td>-0.014</td>
<td>0.371</td>
<td>0.167</td>
<td>0.10±0.17 (±1.700)</td>
</tr>
<tr>
<td>$b_5$</td>
<td>0.015</td>
<td>-0.006</td>
<td>0.014</td>
<td>-0.051</td>
<td>-0.006</td>
<td>0.01±0.03 (±3.000)</td>
</tr>
<tr>
<td>during HD-tDCS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameter</td>
<td>Subject 1</td>
<td>Subject 2</td>
<td>Subject 3</td>
<td>Subject 4</td>
<td>Subject 5</td>
<td>Mean±Std. Dev. (CV = Std. Dev/mean)</td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSE</td>
<td>0.041</td>
<td>0.067</td>
<td>0.089</td>
<td>0.162</td>
<td>0.166</td>
<td>0.11±0.06 (±0.546)</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>--------------------</td>
</tr>
</tbody>
</table>