

Time-varying effective EEG source connectivity: the optimization of model parameters *

M. Rubega, D. Pascucci, J. Rué Queralt, P. Van Mierlo, P. Hagmann, G. Plomp, and C.M Michel

Abstract— Adaptive estimation methods based on general Kalman filter are powerful tools to investigate brain networks dynamics given the non-stationary nature of neural signals. These methods rely on two parameters, the model order p and adaptation constant c , which determine the resolution and smoothness of the time-varying multivariate autoregressive estimates. A sub-optimal filtering may present consistent biases in the frequency domain and temporal distortions, leading to fallacious interpretations. Thus, the performance of these methods heavily depends on the accurate choice of these two parameters in the filter design. In this work, we sought to define an objective criterion for the optimal choice of these parameters. Since residual- and information-based criteria are not guaranteed to reach an absolute minimum, we propose to study the partial derivatives of these functions to guide the choice of p and c . To validate the performance of our method, we used a dataset of human visual evoked potentials during face perception where the generation and propagation of information in the brain is well understood and a set of simulated data where the ground truth is available.

I. INTRODUCTION

In the field of neuroscience, the interest in studying the dynamical causal interactions that characterize rest- or task-related brain networks has led to a remarkable increase in the application of adaptive estimation algorithms. In particular, Granger causality based on adaptive filtering algorithms represent a suitable approach to study dynamical networks composed by highly non-stationary neural signals (e.g., electroencephalogram (EEG) signals) [1] [2]. While adaptive filtering allows dealing with time-varying multivariate time-series, the interpretation and test of direct causal influences strictly depends upon the optimal choice of parameters in the filter settings. A signal x Granger-causes another signal y if the history of x contains information that helps to predict y above and beyond the information contained in the history of y alone [3]. Such dependency on the history of multiple signals implies a model that should take into account the optimal lag (model order) at which causal influences occur and the rapidity with which they evolve. In time-varying multivariate autoregressive (tv-MVAR) models based on the general Kalman filter [4], these two aspects need to be taken into

account as a mandatory step during the filter design: a suboptimal filter that is unable to track the true dynamics of a system can easily increase the risk of wrong causality inference. For this reason, it is imperative to rely on objective criteria for the selection of the two key Kalman parameters that determine the structure and speed of the estimated causal influences. The model order p indicates the number of lags to be considered during modeling, and the adaptation constant c is the tuning parameter that controls the tracking speed and smoothness of the tv-MVAR estimates [5] [6].

The model order selection in tv-MVAR models is not trivial [7] and has been usually performed by means of information-based criteria that neglect the non-stationary nature of the signals (e.g., Akaike information criterion, AIC [8], and Bayesian information criterion, BIC [9]) [10] [11] or through modified time-varying versions of the same criteria (e.g., the modified AIC, MAIC [12]). Other approaches have relied on choosing the model order a priori [13] or based on a comparison between parametric and non-parametric spectral density estimates [14] [15], thus, considering only the univariate part of the tv-MVAR system.

The optimal adaptation constant has been also determined via similar a priori and information-based approaches [15] [16] or through different cost functions that minimize residual errors [10] [17] which are prone to the risk of overfitting real data. In more recent years, a new method has been proposed to estimate the adaptation constant c , based on the predictive least square principle (PLS) [5]. Alternative, to overcome the use of adaptation constants, windowing approaches have also been proposed [18] [19], but with some critical limitations: high temporal resolution requires short time windows which lead to few residuals when assessing the quality of the fit, the model order also depends on the size of the window and both the choice of the size of the window and the overlap percentage are subjective.

Given the lack of a standard selection procedure and the tendency of residual- and information-based criteria to promote overfitting in which an absolute minimum is not guaranteed, in the present work we test a new method based on the partial derivatives of a residual minimization function.

* Research supported by the Swiss National Science Foundation: Grant No. CRSII5_170873, and Grant No. 320030_159705, by the National Centre of Competence in Research (NCCR) "SYNAPSY—The Synaptic Basis of Mental Diseases": Grant No. 51NF40-158776, and by the European Union's Horizon 2020 research and innovation program: Grant No. 660230.

M. R. is with the Functional Brain Mapping Lab, Department of Fundamental Neurosciences, University of Geneva, Geneva, 1202 Switzerland (corresponding author to provide phone: +41 0223790865; e-mail: maria.rubega@unige.ch).

D. P. is with the Perceptual Networks Group, Department of Psychology, University of Fribourg, Fribourg, 1700 Switzerland (e-mail: davis.pascucci@unifr.ch).

J. R. Q. is with the Perceptual Networks Group, Department of Psychology, University of Fribourg, Fribourg, 1700 Switzerland and with the

Department of Radiology, University Hospital of Lausanne, Lausanne, Switzerland (e-mail: joan.ruequeralt@unifr.ch).

P. V. M is with the Medical Image and Signal Processing Group, Department of Electronics and Information Systems, Ghent University, 9000 Ghent, Belgium (e-mail: pieter.vanmierlo@ugent.be).

P. H. is with the Department of Radiology, University Hospital of Lausanne, Lausanne, Switzerland (e-mail: patric.hagmann@epfl.ch).

G. P. is with the Perceptual Networks Group, Department of Psychology, University of Fribourg, Fribourg, 1700 Switzerland (e-mail: gijs.plomp@unifr.ch).

C. M. M. is with the Functional Brain Mapping Lab, Department of Fundamental Neurosciences, University of Geneva, Geneva, 1202 Switzerland and with Lemanic Biomedical Imaging Centre (CIBM), 1015 Lausanne, Switzerland (e-mail: christoph.michel@unige.ch).

This approach allows the determination of the inflection point, i.e., a point on a continuously differentiable plane curve at which the function changes from being concave to convex, which corresponds to our optimal choice of p and c . Because of the lack of an objective ground truth, the evaluation of the performance of effective connectivity methods on real data sets is challenging. As first step to demonstrate the validity of the proposed method, we used a dataset of visual evoked potentials (VEP) during face perception where the generation and propagation of information in the brain is relatively well understood. Complementarily, we tested the method in a set of simulated data imposing p and c a priori.

II. DATA

A. Visual Evoked Potential of face perception

Many behavioral studies have investigated the process involved in visual stimuli such as face images [20] [21]. Traditional measures are based on the N170 face-sensitive evoked response component [22]. Human faces evoke a large negative potential (N170) over the occipital-parietal scalp, more prominent over the right than the left hemisphere, which is reduced in evoked potentials elicited by other animate and inanimate non-face stimuli [23]. Applying effective connectivity in face perception, i.e., describing the network of directional effects of one brain region over another, may be a powerful instrument to study this visual process. In order to study these causal effects, it is important to precisely reconstruct the face-response stimulus in the source space. The accuracy of the model reconstruction depends on the choice of the model parameters. We expected that our tv-MVAR estimates would reliably identify the main components of the VEP (i.e., the P100 and the N170 in the source space), and that the major driver of the network would be localized in the lateral, basal temporal and occipital cortices including the fusiform gyrus as shown through EEG changes from implanted electrodes [24] [25] [26].

B. Experimental protocol

Participants (13 subjects, 2 males, age=24.15 \pm 3.41 y) sat in a dimly lit sound-attenuated and electrically shielded room with their head positioned on a chinrest at \sim 70 cm from the monitor. Each trial lasted 1.2 s and started with a blank screen lasting 500 ms. After the blank interval, one image (either a face or a scramble image) was presented for 200 ms and participants had the remaining 1000 ms to respond. The task was to report whether they saw a face or not (Yes/No task) by pressing two buttons on a response box. Faces and scrambled images were randomly interleaved across trials. After the participant's response, there was a random interval (from 600 to 900 ms) before the beginning of a new trial. The experiment consisted of 4 blocks of 150 trials each, for a total of 600 trials, i.e., 300 with faces and 300 with scrambled images. For this study, we used the EEG data in response of the face images (300 trials per subject). During the experiment, EEG data were recorded continuously at 1024 Hz through a 128-channel Biosemi Active Two EEG system (Biosemi, Amsterdam, The Netherlands). Experimental procedures were conformed to Swiss Ethics Committee standards and were approved by the regional ethics committee (CER-VD).

C. Preprocessing

The VEP EEG signals were downsampled at fs=200 Hz and detrended to remove slow fluctuations and linear trends [57]. The line and monitor noise (50 and 75 Hz, plus

harmonics) were attenuated with an adaptive multitaper filter (Cleanline plugin for EEGLAB). EEG epochs were then extracted from the continuous dataset and time-locked from -1000 ms to 1000 ms relative to the onset of each image. Noisy channels were identified by visual inspection and removed before preprocessing. Individual epochs containing non-stereotyped artifacts, peri-stimulus eye blinks and eye movements (occurring within \pm 500 ms from stimulus onset) were also identified by visual inspection and removed from further analysis (mean number of epochs removed across participants: 6 \pm 5). Data were cleaned from remaining physiological artifacts (eye blinks, horizontal and vertical eye movements, muscle potentials and other artifacts) through a PCA-informed ICA algorithm implemented in EEGLAB. After ICA cleaning, the identified artifact channels were interpolated using the nearest-neighbor spline method and the data were re-referenced to the average reference.

D. Simulated Data

To test if the proposed method is capable of detecting the correct values of the model order and the adaptation constant, we generated two sets of simulated data imposing p and c a priori. Seventy-eight networks with 82 nodes were generated through a random discrete state-space model adding a single-trial VEP component and observation noise [27]. The amplitude of the surrogate tv-MVAR coefficients decreased exponentially with the lags following an underdamped oscillator. VEPs were generated perturbing the univariate values of the simulated tv-MVAR matrices. Asymptotical stability of the tv-system was imposed as in [28]. Each simulated network contains 100 trials of 800-sample signals with signal-to-noise ratio (SNR) equal to 1, model order equal to 5, 7 or 9 and the adaptation constant equal to $3 \cdot 10^{-3}$ or $3 \cdot 10^{-2}$.

III. METHOD

In the framework of a MATLAB toolbox (code available upon request to the authors) that implements the adaptive Kalman filtering [4] and information Partial Directed Coherence (iPDC) in the source space [29], we propose a new analytical approach for the optimal choice of the Kalman parameters. (We refer the reader to [30] for further details on the implementation of the effective connectivity pipeline in the source space from high-density EEG recordings). In order to estimate in parallel the adaptation constant and the tv-MVAR model order, we refined the predictive least square principle (PLS) [5] [31]:

$$\widehat{PLS}_{(p,c)} = \frac{1}{N} \sum_{t=1}^N (1-c)^{N-t} \|PE_p(t)\|^2 \quad (1)$$

where N is the number of samples, c is the adaptation constant [4], PE is the true prediction error function of t and the model order p [32]). $\widehat{PLS}(p, c)$ is a method based on model residuals and should be minimized to choose p and c . However, as for other residual- and information-based criteria in the context of tv-MVAR models (AIC, BIC), the cost function is not ensured to reach its absolute minimum and tends to monotonically decrease with overfitting (Fig. 1a). This could lead to too high model order and adaptation constant that would increase both computational cost and noise sensitivity. If an absolute minimum does not exist, we propose to study the partial derivative of the cost function $\widehat{PLS}(p, c)$. Our aim was to find the value at which the PE does not significantly improve by increasing p and c . Thus, we

estimated the inflection point of the cost function $\widehat{PLS}(p, c)$ studying its partial derivatives $\frac{\partial(\cdot)}{\partial p}$ along p coordinate axis. We kept the variable c fixed and differentiated the resulting function as a function of variable p . Geometrically, the partial derivative gives the slope of the function at each point in the directions parallel to the p coordinate axis. We estimated the negative values of $\frac{\partial^3}{\partial p^3} \widehat{PLS}(p, c)$ and we chose (p_{opt}, c_{opt}) such that both $\frac{\partial}{\partial p} \widehat{PLS}(p, c)$ and $\frac{\partial^2}{\partial p^2} \widehat{PLS}(p, c)$ tend to zero in the neighborhood of (p_{opt}, c_{opt}) . These values of p and c indicate our inflection point, i.e., the point of the cost function where the curvature changes its sign while a tangent exists. To test the goodness of the data fitting using the values (p_{opt}, c_{opt}) for each subject, we computed the residuals from the data and the model estimates, the *iPDC* outflow from the tv-MVAR coefficients, and parametric and non-parametric power spectra of VEP dataset.

IV. RESULTS

As a representative example, the cost function $\widehat{PLS}(p, c)$ is reported in Fig. 1a for a representative subject. Given that $\widehat{PLS}(p, c)$ is a differentiable function and the lack of a global minimum, we could study its partial derivatives (Fig. 1 b-c-d) and estimate the inflection point. Considering all subjects, the model order p was estimated in the interval [4:9], i.e., [20: 45] ms, and the adaptation constant c spanned between [1.6: 5.4] 10^{-3} . (In parallel, we also exploited the time-varying formulation of AIC and BIC as cost function varying p and c in the same intervals, the 96% of the estimated values tended to $-\infty$, rendering a reliable estimation of the absolute minimum and the computation of the partial derivatives impossible.) The obtained values for the model order p satisfied the hypothesis on the average maximum human brain fiber length (circa 15 cm) and conduction delay (circa 6 m/s) [33]. Whereas, the values for the adaptation constant c are in the same interval to the one suggested by Schögl in [6] and exploited in [4]. Fig. 2 reports a comparison between the model prediction (Fig. 2b) and the original data (Fig. 2a) for a representative subject. It is clear that there is a strong correlation between the model prediction and the original data, $r > 0.5$ (Fig. 2c). Fig. 3 depicts the average across subject of the power in [1-40] Hz for our 82 regions of interest (ROIs). The power was computed both directly from the reconstructed source waveforms by a non-parametric approach (Fig. 3a) and from the tv-MVAR coefficients (Fig. 3b) during N170. We found a strong correlation between the distributions of power in the brain regions computed with the two different methods. Indeed, ROIs with the maximal power (>95% percentile) were localized in the right lateral occipital cortex and in the right inferior temporal cortex with both methods. Finding the regions with highest power with a non-parametric approach might already give enough information about the main drivers in a network, but we should take into account that non-parametric approaches are less robust to artifacts and noise in the EEG than parametric ones.

Exploiting the tv-MVAR coefficients, we computed the $|iPDC|$ values during the first 500 ms after the stimulus, and we compared the values of the outflow from each ROIs at N170. The connectivity patterns between the different cortical regions were summarized by representing the total outflow from a cortical region toward the others, generated by the sum

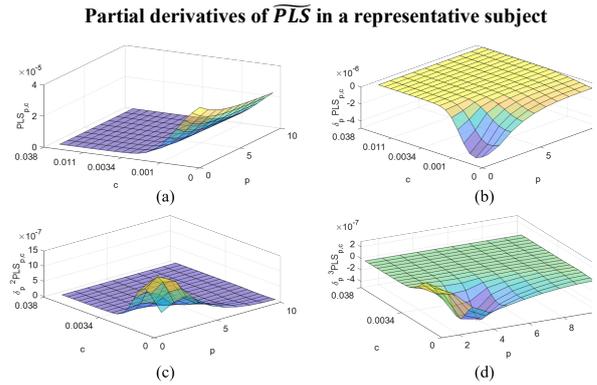


Figure 1. Values of (a) $\widehat{PLS}_{p,c}$ and its (b) first-, (c) second- and (d) third-partial derivative along p axis in a representative subject.

Source waveform in the left inferior temporal cortex in a representative subject

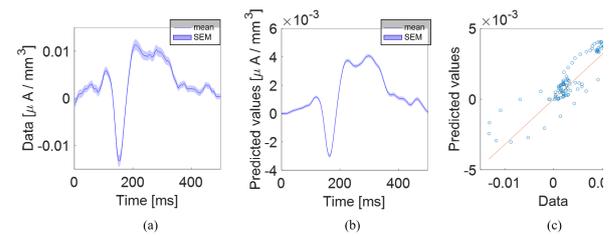


Figure 2. (a) Reconstructed source waveform in the left inferior temporal cortex in a representative subject. (b) Predicted values by the tv-MVAR model ($p = 4$; $c = 0.0043$) for the source waveform in the same brain region of (a). (c) Scatter plot of the original data in (a) versus the predicted values in (b).

Non-parametric and parametric average power in all subjects

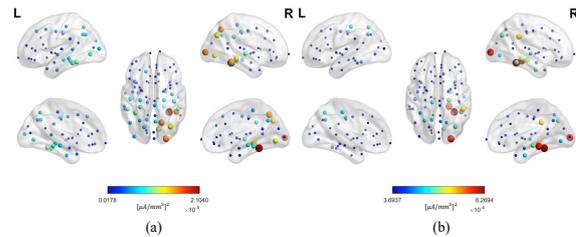


Figure 3. (a) Non-parametric and (b) parametric average power [1-40] Hz across subjects computed from the source waveforms in the 82 ROIs during N170. Nodes dimension and color identify the value of the power.

Average outflow in all subjects

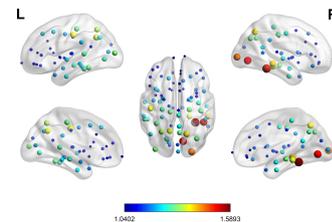


Figure 4. Average outflow across subjects in the 82 ROIs computed from *iPDC* matrix during N170. Nodes dimension and color identify the value of the outflow.

of all the statistically significant links obtained by application of the *iPDC* to the cortical waveforms (with their values). In Fig. 4, the total outflow during N170 for each ROI is represented by a sphere centered on the cortical region, whose radius is linearly related to the magnitude of all the out-coming directed links to the other regions. Such information is also coded through a color scale. The greatest amount of

information outflow depicts the ROI as one of the main sources (drivers) of functional connections to the other ROIs [34]. The ROIs with the maximum outflow (> 95% percentile) were localized in the right lateral-occipital cortex, and in the inferior temporal cortex (Fig. 4). Our findings are in accordance with the ones in the literature. N170 usually displays right-hemisphere lateralization and its response in the scalp is maximal over occipito-temporal electrodes. Regarding the source of N170, this event-related potential was shown to be generated in widely distributed occipital-temporal-parietal cortical regions.

In Tab. 1 the results of the set of simulations are reported: the 2nd column stands for the p and c imposed a priori; the 3rd and the 4th columns to c predicted and to the p predicted (mean±std) by 100 trials of each simulated network. Both the estimated values for the adaptation constant and the model order satisfied the hypothesis.

TABLE I. SIMULATION RESULTS

# simulations x (#trials, # nodes, SNR, # samples)	Model parameters		
	(p, c)	$c_{predicted}$	$p_{predicted}$
13 x (100, 82, 1, 800)	($5, 3 \cdot 10^{-3}$)	$[1 \pm 0.3] \cdot 10^{-3}$	5 ± 2
	($7, 3 \cdot 10^{-3}$)	$[1 \pm 6] \cdot 10^{-3}$	8 ± 2
	($9, 3 \cdot 10^{-3}$)	$[1 \pm 0.1] \cdot 10^{-3}$	9 ± 3
	($5, 3 \cdot 10^{-2}$)	$[1 \pm 0.5] \cdot 10^{-2}$	4 ± 2
	($7, 3 \cdot 10^{-2}$)	$[1 \pm 0.5] \cdot 10^{-2}$	7 ± 4
	($9, 3 \cdot 10^{-2}$)	$[3 \pm 2] \cdot 10^{-2}$	9 ± 3

V. CONCLUSION

In this work, we proposed to study the partial derivatives of residual- and information-based criteria to guide the choice of p and c when these functions do not reach an absolute minimum. We reported the results in both surrogate and real data testing the partial derivatives of \widetilde{PLS} . The consistence of the results of the simulations and in finding the main drivers during N170 prove the validity of adaptive estimation algorithms in estimating tv-Granger causality after an accurate choice of the model order and adaptation constant.

This is an important first step towards automated model selection in tv-MVAR modeling.

REFERENCES

- [1] E. Pereda, R. Q. Quiroga and J. Bhattacharya, "Nonlinear multivariate analysis of neurophysiological signals," *Progress in neurobiology*, vol. 77, no. 1-2, pp. 1-37, 2005.
- [2] A. K. Seth, A. B. Barrett and L. Barnett, "Granger causality analysis in neuroscience and neuroimaging," *Journal of Neuroscience*, vol. 35, no. 8, pp. 3293-3297, 2015.
- [3] C. W. J. Granger, "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica*, vol. 37, no. 3, pp. 424-438, 1969.
- [4] T. Milde, L. Leistriz, L. Astolfi, H. R. Miltner, T. Weiss, F. Babiloni and H. Witte, "A new Kalman filter approach for the estimation of high-dimensional time-variant multivariate AR models and its application in analysis of laser-evoked brain potentials," *NeuroImage*, vol. 50, pp. 960-969, 2010.
- [5] A. H. Costa and S. Hengstler, "Adaptive time-frequency analysis based on autoregressive modeling," *Signal Processing*, vol. 91, pp. 740-749, 2011.
- [6] A. Schlögl, The electroencephalogram and the adaptive autoregressive model: theory and applications, Germany: Shaker, 2000.
- [7] A. H. Omidvarnia, M. Mesbah, M. S. Khlif, J. M. O'Toole, P. B. Colditz and B. Boashash, "Kalman filter-based time-varying cortical connectivity analysis of newborn EEG," in *Engineering in Medicine and Biology Society (EMBC)*, 2011.
- [8] H. Akaike, "Maximum likelihood identification of Gaussian autoregressive moving average models," *Biometrika*, vol. 60, no. 2, pp. 255-265, 1973.
- [9] G. Schwarz, "Estimating the dimension of a model," *The annals of statistics*, vol. 6, no. 2, pp. 461-464, 1978.
- [10] M. Arnold, X. H. R. Milne, H. Witte, R. Bauer and C. Braun, "Adaptive AR modeling of nonstationary time series by means of Kalman filtering," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 5, pp. 553-562, 1998.
- [11] D. Pascucci, A. Hervais-Adelman and G. Plomp, "Gating by induced A-G asynchrony in selective attention.," *Human Brain Mapping*, vol. 39, no. 10, pp. 3854-3870, 2018.
- [12] S. Goto, M. Nakamura and K. Uosaki, "On-line spectral estimation of nonstationary time series based on AR model parameter estimation and order selection with a forgetting factor," *IEEE Transactions on Signal Processing*, vol. 43, no. 6, pp. 1519-1522, 1995.
- [13] A. Atiyabi, F. Shic and A. Naples, "Mixture of autoregressive modeling orders and its implication on single trial EEG classification," *Expert systems with applications*, vol. 65, pp. 164-180, 2016.
- [14] C. Porcaro, F. Zappasodi, P. M. Rossini and F. Tecchio, "Choice of multivariate autoregressive model order affecting real network functional connectivity estimate," *Clinical Neurophysiology*, vol. 120, no. 2, pp. 436-448, 2009.
- [15] E. G. Ghumare, M. Schrooten, R. Vandenberghe and P. Dupont, "A Time-Varying Connectivity Analysis from Distributed EEG Sources: A Simulation Study.," *Brain topography*, pp. 1-17, 2018.
- [16] E. Ghumare, M. Schrooten, R. Vandenberghe and P. Dupont, "Comparison of different Kalman filter approaches in deriving time varying connectivity from EEG data," in *Engineering in Medicine and Biology Society (EMBC)*, 2015.
- [17] A. Schlogl, J. S. Roberts and G. Pfurtscheller, "A criterion for adaptive autoregressive models," in *Engineering in Medicine and Biology Society (EMBC)*, 2000.
- [18] M. Ding, S. L. Bressler, W. Yang and H. Liang, "Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment," *Biological cybernetics*, vol. 83, no. 1, pp. 35-45, 2000.
- [19] G. M. Hoerzer, S. Liebe, A. Schloegl, N. K. Logothetis and G. Rainer, "Directed coupling in local field potentials of macaque v4 during visual short-term memory revealed by multivariate autoregressive models," *Frontiers in computational neuroscience*, vol. 4, p. 14, 2010.
- [20] J. V. Haxby, E. A. Hoffman and M. I. Gobbini, "The distributed human system for face perception," *Trends in cognitive sciences*, vol. 4, no. 6, pp. 223-233, 2000.
- [21] E. Hoffman and J. Haxby, "Distinct representations of eye gaze and identity in the distributed human neural system for face perception," *Nature Neuroscience*, vol. 10, no. 3, pp. 80-84, 2000.
- [22] B. Rossion and S. Caharel, "ERP evidence for the speed of face categorization in the human brain: Disentangling the contribution of low-level visual cues from face perception," *Vision research*, vol. 51, pp. 1297-1311, 2011.
- [23] S. Bentin, T. Allison, A. Puce, E. Perez and G. McCarthy, "Electrophysiological studies of face perception in humans," *Journal of cognitive neuroscience*, vol. 8, no. 6, pp. 551-565, 1996.
- [24] K. J. Miller, D. Hermes, F. Pestilli, G. S. Wig and J. G. Ojemann, "Face percept formation in human ventral temporal cortex.," *Journal of Neurophysiology*, vol. 118, no. 5, pp. 2614-2627, 2017.
- [25] F. Baroni, J. van Kempen, H. Kawasaki, C. K. Kovach, H. Oya, M. A. Howard, R. Adolphs and N. Tsuchiya, "Intracranial markers of conscious face perception in humans," *Neuroimage*, vol. 162, pp. 322-343, 2017.
- [26] C. M. Hamamé, J. R. Vidal, M. Perrone-Bertolotti, T. Ossandon, K. Jerbi, P. Kahane, O. Bertrand and J. P. Lachaux, "Functional selectivity in the human occipitotemporal cortex during natural vision: evidence from combined intracranial EEG and eye-tracking," *Neuroimage*, vol. 95, pp. 276-286, 2014.
- [27] P. L. Rodrigues and L. A. Baccalá, "A new algorithm for neural connectivity estimation of EEG event related potentials," in *Engineering in Medicine and Biology Society (EMBC)*, 2015.
- [28] A.N. Vargas, J. BR. Do Val, E. F. Costa, "On stability of linear time-varying stochastic discrete-time systems.," in *European Control Conference (ECC)*. IEEE, p. 2423-2427, 2007.
- [29] D. Y. Takahashi, L. A. Baccala and K. Sameshima, "Information theoretic interpretation of frequency domain connectivity measures," *Biological cybernetics*, vol. 103, pp. 463-469, 2010.
- [30] M. Rubega, M. Carboni, M. Seeber, D. Pascucci, S. Tourbier, G. Toscano, P. Van Mierlo, P. Hagmann, G. Plomp, S. Vulliemoz and C. M. Michel, "Estimating EEG Source Dipole Orientation Based on Singular-value Decomposition for Connectivity Analysis," *Brain Topography*, pp. 1-16, 2018.
- [31] J. Rissanen, "A predictive least-squares principle," *IMA Journal of Mathematical Control and Information*, vol. 3, no. 2.3, pp. 211-222, 1986.
- [32] M. Wax, "Order selection for AR models by predictive least squares," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 581-588, 1988.
- [33] R. Caminiti, F. Carducci, C. Piervincenzi, A. Battaglia-Mayer, G. Confalone, F. Visco-Comandini, P. Pantano and G. Innocenti, "Diameter, length, speed, and conduction delay of callosal axons in macaque monkeys and humans: comparing data from histology and magnetic resonance imaging diffusion tractography.," *Journal of Neuroscience*, vol. 33, no. 36, pp. 14501-14511, 2013.
- [34] F. Babiloni, F. Cincotti, C. Carducci, C. Babiloni, F. Carducci, D. Mattia, L. Astolfi, A. Basilisco, P. M. Rossini, L. Ding, Y. Ni, J. Cheng, K. Christine, J. Sweeney and B. He, "Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function.," *NeuroImage*, vol. 24, no. 1, pp. 118-131, 2005.