Optimizing spatial filters for the extraction of envelope-coupled neural oscillations

Sven Dähne*, Vadim Nikulin†, David Ramírez‡, Peter J. Schreier‡, Klaus-Robert Müller*¶, and Stefan Haufe§*

*Dept. of Machine Learning, Berlin Institute of Technology, Berlin, Germany

†Dept. of Neurology, Charité University Medicine Berlin, Berlin, Germany

‡Dept. of Electrical Engineering and Information Technology, Universität Paderborn, Paderborn, Germany

§Dept. of Biomedical Engineering, The City College of New York, New York City, USA

¶Dept. of Brain and Cognitive Engineering, Korea University, Seoul, Korea

***Abstract*—Amplitude-to-amplitude interactions between neural oscillations are of a special interest as they show how the strength of spatial synchronization in different neuronal populations relates to each other during a given task. While, previously, amplitude-to-amplitude correlations were studied primarily on the sensor level, we present a source separation approach using spatial filters which maximize the correlation between the envelopes of brain oscillations recorded with electro-/magnetencephalography (EEG/MEG) or intracranial multichannel recordings. Our approach, which is called canonical source power correlation analysis (cSPoC), is thereby capable of extracting genuine brain oscillations solely based on their assumed coupling behavior even when the signal-to-noise ratio of the signals is low.**

I. Introduction

The investigation of neural rhythms is a major aspect of the ongoing endeavor to understand the principles of brain functions [1]. In the past decades, research on brain oscillations measured by electrophysiological (e. g., electro-/magnetencephalography, EEG/MEG, and electrocorticography, ECoG) recordings has provided a large body of evidence suggesting a key role of neural rhythms in sensory processing [2].

Interactions between oscillatory processes have been of particular interest in the recent past [3]–[5]. Different types of cross-frequency coupling between neural systems have been identified and associated with specific brain functions. Here we focus on power-to-power (or amplitude-to-amplitude) coupling [6], which describes the interaction between the spectral power (or the envelope/amplitude) of distinct oscillations at specific frequencies or narrow frequency bands.

The problem of recovering the underlying neural signals from multivariate sensor readings is called (blind) source separation (BSS), and is inherently ill-posed, meaning that it relies on the presence of prior knowledge about these signals. The more precise the prior knowledge is, the better the true underlying brain processes can be recovered. For example, Independent Component Analysis (ICA) [7] assumes mutual statistical independence of the sources, and solves the BSS problem optimally if this assumption is met. If one is, however, interested in recovering sources exhibiting a specific type of dependency (among each other, or to an external variable) – such as power-to-power coupling of brain oscillations – it is reasonable to base the reconstruction of the source activity exactly on this assumed dependency rather than on mutual independence.

978-1-4799-4149-0/14/$31.00 ©2014 IEEE

Source separation with respect to the power dynamics of the sources has been addressed before. Recent approaches have focused on the extraction of neuronal oscillations with band-power dynamics that are correlated with either univariate target signals such as stimulus- or cognitive variables [8], or with multivariate target signals such as signals from a different measurement modality [9]. The main idea of the mentioned approaches is to i) *directly optimize the quantity of interest*, namely the correlation between the band-power time course and the time course of a potentially related signal, and ii) adhere to the linear generative model underlying electrophysiological recordings. Thereby the mixing of local brain activity that is due to volume conduction in the head can be reversed and local brain oscillations can be recovered with high signal-to-noise ratio (SNR).

In this contribution we present a novel unsupervised analysis method for the optimal extraction of pairs of neuronal sources the band-power dynamics of which are correlated with each other. We refer to this method as *canonical Source Power Correlation* analysis (cSPoC) and we demonstrate its utility by way of extensive and realistic simulation scenarios. We evaluate the robustness of cSPoC with respect to signal-to-noise ratio (SNR) as well as to strength of true correlations present in the simulated data and compare its performance to alternative state-of-the-art methods.

II. Methods

A. Forward Model

Let the measured data be represented by the multivariate variable $\mathbf{x}(t) \in \mathbb{R}^N$, where N denotes the number of sensors. Moreover, let $\mathbf{s}(t) \in \mathbb{R}^K$ denote the time courses of the electrical source activity of some underlying neural processes of interest. In electrophysiological recordings, the physics of volume conduction implies a linear mapping from these neural sources to the measurement [10]–[12]. The generative (or *forward*) model thus reads

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \epsilon(t) \;, \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{N\times K}$ contains mixing coefficients describing the influence of the source activity on each sensor and $\epsilon(t) \in \mathbb{R}^N$ models noise contributions from outside the brain. The columns of $\mathbf{A} = \left(\mathbf{a}^1, \ldots, \mathbf{a}^K\right)$ are referred to as spatial *activation patterns* of the respective neural sources. Thus, $\mathbf{a}^i$, i.e. the i$^{\text{th}}$ column of $\mathbf{A}$, is the electrical potential or magnetic field pattern of the i$^{\text{th}}$ neural source.

B. Backward Model

We are ultimately interested in the time-courses (rows of $\mathbf{s}$) and corresponding activation patterns (columns of $\mathbf{A}$) of power-coupled neural sources in the datasets. However, estimating both the sources and their activation patterns jointly leads to difficult optimization problems. We can reduce the computational complexity by resorting to a so-called discriminative (or *backward*) modeling approach. Here, the time-courses of P neural sources are estimated by projecting the data linearly onto a set of spatial *extraction filters* $\mathbf{W} \in \mathbb{R}^{N\times P}$:

$$\hat{\mathbf{s}} = \mathbf{W}^\top \mathbf{x}(t) \;. \tag{2}$$

It can be shown that for each backward model of the form Eq. (2), a *corresponding* forward model of the form Eq. (1) exists, for which the noise $\epsilon(t)$ is uncorrelated to the sources $\hat{\mathbf{s}}(t)$ [12]. The activation patterns of this forward model can be obtained from the spatial filters via the transformation

$$\mathbf{A} = \mathbf{C}_\mathbf{x}\mathbf{W}\mathbf{C}_{\hat{\mathbf{s}}}^{-1} \;, \tag{3}$$

where $\mathbf{C}_\mathbf{x}$ denotes the covariance matrix of the data $\mathbf{x}(t)$, and $\mathbf{C}_{\hat{\mathbf{s}}}$ is the covariance matrix of the extracted sources $\hat{\mathbf{s}}(t)$ [12].

C. Canonical source power correlation analysis (cSPoC)

Here we want to find pairs of oscillatory neural sources from electrophysiological measurements based on the assumption that the envelope dynamics of these sources are correlated. Thus, we require the presence of two sets of measured data represented by the multivariate variables $\mathbf{x}$ and $\mathbf{x}'$, which are assumed to be generated according to Eq. (1). Furthermore, the datasets $\mathbf{x}$ and $\mathbf{x}'$ are assumed to be related by common source processes among the rows of $\mathbf{s}_x$ and $\mathbf{s}_{\acute{x}}$, whose envelopes are linearly correlated.

Let $\mathbf{w}_\mathbf{x} \in \mathbb{R}^{N_x}$ and $\mathbf{w}_{\mathbf{x}'} \in \mathbb{R}^{N_{\acute{x}}}$ be extraction filters for a single source pair $s_\mathbf{x} = \mathbf{w}_\mathbf{x}^\top \mathbf{x}$ and $s_{\mathbf{x}'} = \mathbf{w}_{\mathbf{x}'}^\top \mathbf{x}'$. The instantaneous amplitudes (envelopes) of these sources are given by

$$\Phi = \sqrt{(\mathbf{w}^\top \mathbf{x})^2 + (\mathbf{w}^\top \mathrm{H}[\mathbf{x}])^2}, \tag{4}$$

where $\mathrm{H}[\cdot]$ denotes the Hilbert transform. Accordingly for $\mathbf{s}_{\mathbf{x}'}$.

With these definitions, we can expresses the correlation between the source envelopes as a function of the spatial filters $\mathbf{w}_\mathbf{x}$ and $\mathbf{w}_{\mathbf{x}'}$. The cSPoC objective function thus reads

$$\max_{\mathbf{w}_\mathbf{x}, \mathbf{w}_{\mathbf{x}'}} \; \mathrm{Corr}\left(\Phi_\mathbf{x}, \Phi_{\mathbf{x}'}\right) , \tag{5}$$

where $\Phi_\mathbf{x}$ and $\Phi_{\mathbf{x}'}$ are defined according to Eq. (4) for $\mathbf{w}_\mathbf{x}$ and $\mathbf{w}_{\mathbf{x}'}$. respectively.

The objective stated in Eq. (5) is a non-convex higher-order nonlinear function of the coefficients $\mathbf{w}_\mathbf{x}$ and $\mathbf{w}_\mathbf{y}$. We propose to optimize it by means of standard nonlinear optimization techniques such as the limited-memory Broyden-Fletcher-Goldfarb-Shanno (l-BFGS) algorithm [13] implemented in MATLAB's (The Mathworks) `fminunc` routine.

D. Simulations

The following equations describe the generation of pseudo-EEG measurements $\mathbf{x}(t)$ as the sum of contributions $\mathbf{x}_\text{s}(t)$ due a single target source, contributions $\mathbf{x}_\text{ba}(t)$ due to background sources, and sensor noise $\mathbf{x}_\text{noise}(t)$. Pseudo-EEG measurements $\mathbf{x}'$ are generated in the same way.

$$\mathbf{x}_\text{s}(t) = \mathbf{a}_\mathbf{x}^1 \left(\mathbf{s}_x(t)\right)_1 \tag{6}$$

$$\mathbf{x}_\text{ba}(t) = \sum_{i=2}^{31} \mathbf{a}_\mathbf{x}^i \left(\mathbf{s}_x(t)\right)_i \tag{7}$$

$$\mathbf{x}_\text{noise}(t) = \frac{1}{c_\text{ba}}\, \mathbf{x}_\text{ba}(t) + \gamma_\epsilon\, \epsilon_\mathbf{x}(t) \tag{8}$$

$$\mathbf{x}(t) = \gamma \frac{\mathbf{x}_\text{s}(t)}{c_\text{s}} + \frac{\mathbf{x}_\text{noise}(t)}{c_\text{noise}} \tag{9}$$

The constants c_ba, c_s, and c_noise are the (Frobenius-) norm constants of the corresponding data matrices $\mathbf{x}_\text{ba}$, $\mathbf{x}_\text{s}$, and $\mathbf{x}_\text{noise}$, and serve to balance the data matrices prior to building the weighted sums.

The scaling parameter γ_ϵ determines the strength of (Frobenius-) normalized Gaussian noise $\epsilon_\mathbf{x}$, and was set to 0.1 in all simulations. The scaling parameter $\gamma > 0$ determines the ratio between the energy of the scalp-projected target source time courses and the energy of all other (background and sensor noise) contributions in the data. Thus the signal-to-noise ratio in dB scale is given by $\mathrm{SNR} = 10 \log_{10}(\gamma)$.

The number of simulated EEG channels was $N_x = N_{\acute{x}} = 24$ and a total of 31 sources were simulated for each data set. The spatial patterns of the simulated sources were obtained as the scalp potentials induced by placing electrical dipoles at random locations and with random orientations in 3D voxel space. These potentials were calculated within a realistic head model.

The time-courses of the sources $\mathbf{s}_x$ and $\mathbf{s}_{\acute{x}}$ were constructed by specifying an amplitude spectrum, in which the frequency bins of the alpha range (8 to 12 Hz) had non-zero values, while all other frequency bins were set to zero. The corresponding phase spectrum was drawn from a uniform distribution over the interval 0 to 2π and the inverse Fourier transform was applied in order to obtain band-limited oscillations with random envelope modulations.

In each simulation, a total of 10 minutes of data was simulated. The first half of the data was used to train the algorithms, i.e. to optimize a pair of spatial filters $\mathbf{w}_\mathbf{x}$ and $\mathbf{w}_{\mathbf{x}'}$. The second half was then used for testing. As a performance metric we report the correlations between the envelopes of the extracted $\mathbf{x}$ and $\mathbf{x}'$, computed on the test data.

We systematically varied the SNR between target and background sources, as well as the true correlation r_{true} between envelopes of the target sources in the two datasets $\mathbf{x}$ and $\mathbf{x}'$. The simulations were repeated 100 times for each parameter setting of SNR and r_{true}, with each time newly generated data.

We benchmarked cSPoC against two other unsupervised methods, namely CCA and PowerCCA [14]. For CCA, envelopes were computed for each simulated channel and the envelope time courses were used as input. cSPoC and PowerCCA act on time-domain data and envelopes were computed after performing the backward projection, i. e., on the extracted sources. PowerCCA is a recent method that also optimizes spatial filters for two data sets such that the envelopes of the projected signals exhibit maximal correlation.

Since the "ground truth" (i. e., the time-courses and envelopes of the coupled sources) is known in this simulation setting, it is possible to apply *supervised* methods for extracting the sources-of-interest as baseline methods. Here, we applied ordinary least squares regression (OLS) to datasets $\mathbf{x}$ and $\mathbf{x}'$ separately, using the actual time-course of the target sources as a training signal. Since the target source signals are not available in practice, OLS mainly serves as an empirical upper bound for the performance in the benchmarking.

III. Results

The results of the simulations are shown in Figure 1. The figure depicts the correlations (obtained on validation data) between the envelopes of the extracted sources from simulated datasets $\mathbf{x}$ and $\mathbf{x}'$. In Figure 1 **A** and **B**, we show the recovered correlations as a function of the SNR for two different values (0.7 and 1) of the correlation between the envelopes of the simulated target sources, i.e. the "true" correlation. In Figure 1 **C** and **D** we show the recovered correlations as a function of the "true" correlation for two different SNRs (-5 dB and 0 dB).

All methods approach the true correlation between the target sources as the SNR increases above 5 dB. cSPoC attains the best performance among all unsupervised methods and is on par with the supervised OLS over large SNR ranges. Note the large gap in performance between CCA and its competitors. CCA, by design, is able to invert linear mixtures of correlating times series. However, if the linear model (Eq. (1)) holds for the raw data (as it is the case for EEG data), this implies that the envelopes of the sources are *not* linearly mixed in the channel envelopes, making it difficult for CCA to extract the target envelopes. CCA is able to perform well only in high SNR regimes, when the target sources dominate the simulated EEG signal.

All methods eventually fail as the SNR becomes very low (i. e., below -5 dB) and/or the true correlation approaches zero. However, we observe that cSPoC outperforms its unsupervised peers across wide parameter regimes. cSPoC outperforms PowerCCA with statistical significance in the range of -7 dB and 5 dB and for all correlations larger or equal than 0.3 (paired t-test, $p < 0.01$).

IV. Discussion and Conclusion

In this paper, we have a novel unsupervised source separation approach for finding oscillatory sources with envelope correlations. The method was benchmarked using extensive simulations. Its performance, which is competitive to supervised approaches, makes cSPoC the method of choice in scenarios in which supervised methods are not applicable, such as Hyperscanning settings. It thus provides a versatile addition to other multivariate analysis tools for cross-frequency coupling such as cross-frequency decomposition [15] for phase coupling and multimodal source power correlation [9] for phase-to-power coupling, as well as tools for extracting brain activity from electrophysiological recordings based on other types of dependencies [16]–[19].

Our algorithm is called *canonical* source power correlation analysis, because it bears some conceptual similarity with canonical correlation analysis (CCA) [20]. CCA and its variants [21]–[23] are powerful tools for the extraction of correlated components from two multivariate data sets. Both CCA and cSPoC seek linear projections of two data sets in order to maximize a correlation coefficient defined on the projected data (sources). However, in case of CCA, the correlation is defined between the source time courses directly, whereas in cSPoC the correlation is defined between nonlinear features of the sources, namely their envelopes. Importantly, CCA cannot be used to extract power-coupled brain sources from electrophysiological recordings. While it is conceivable to use a two-step approach, i. e., first compute the envelopes of the raw channel data of the two data sets and then apply CCA, such an approach does not lead to an accurate inversion of the generative model Eq. (1). The reason for this is that the two operations (i) source extraction through linear projection and (ii) nonlinear processing (computation of the envelope) do not commute. By design, cSPoC computes envelopes on the source level *after* projecting the data onto the extraction filters, and thus performs both operation in the order implied by the generative model.

Being a multivariate technique, cSPoC integrates information from all recording channels and thereby achieves much higher SNRs compared to single channel analysis, which leads to stronger effects. By projecting the data into the cSPoC source space, the dimensionality is reduced and thus the problem of channel-wise comparison/correlations (and multiple testing) is avoided.

We believe that cSPoC will be a valuable tool in the quest for understanding the mechanisms and functions of power-to-power coupling in electrophysiological recordings.

Acknowledgment

SD acknowledges funding by the German Research Foundation (DFG) grant no. MU 987/19-1 and support by the Bernstein Center for Computational Neuroscience, Berlin through the graduate program GRK 1589/1. VN acknowledges funding by the German Research Foundation (DFG) grant no. KFO 247 and by the Bernstein Center for Computational Neuroscience, Berlin. PJS acknowledges funding by the Alfried Krupp von

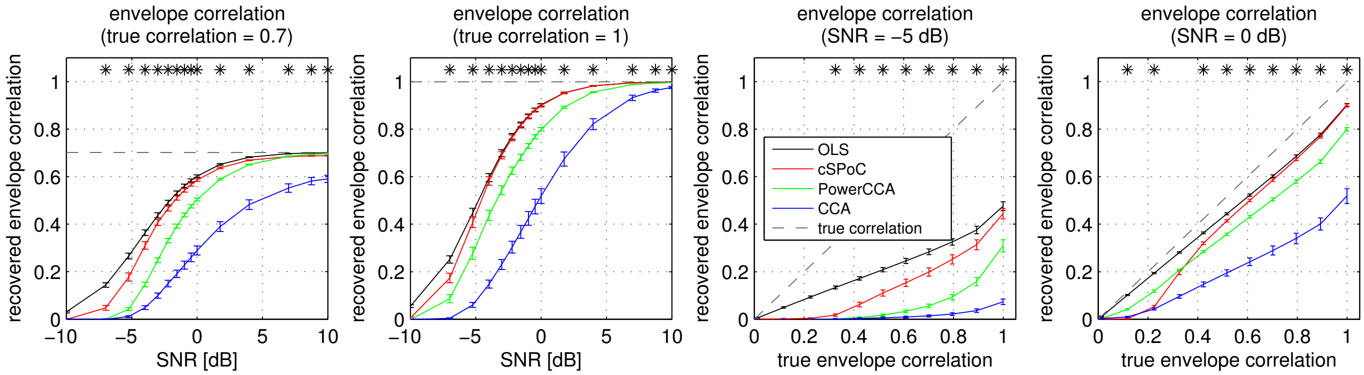

Fig. 1. **Simulation results.** Empirical correlations between envelopes of source pairs extracted by three unsupervised methods (cSPoC, CCA, PowerCCA) and one supervised method (OLS). OLS makes use of the ground truth and is therefore not applicable in practice. Plots **A** and **B**: variation of the SNR for two fixed correlations (0.7 and 1.0, respectively) of the envelopes of the simulated sources ("true" correlations). Plots **C** and **D**: variation of the "true" correlation for fixed SNRs (-5 dB and 0 dB, respectively). Empirical correlations were obtained on validation data, which was not used to fit the models. For each parameter setting, simulations were repeated 100 times with each time newly generated model fitting and validation data. Plotted correlations are averaged across repetitions, and errorbars indicate standard error of the mean (SEM). Black asterisks indicate statistically significant differences between cSPoC and PowerCCA (paired t-test, $p < 0.01$.)

Bohlen und Halbach foundation under its program Return of German scientists from abroad. KRM acknowledges support by the Brain Korea 21 Plus Program through the National Research Foundation of Korea funded by the Ministry of Education. SH acknowledges support by the German Federal Ministry of Education and Research (BMBF), grant No. 01GQ0850.

REFERENCES

[1] G. Buzsáki and A. Draguhn, "Neuronal oscillations in cortical networks." *Science (New York, N.Y.)*, vol. 304, no. 5679, pp. 1926–9, 2004.

[2] G. Thut, A. Nietzel, S. A. Brandt, and A. Pascual-Leone, "Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection." *The Journal of Neuroscience*, vol. 26, no. 37, pp. 9494–9502, 2006.

[3] O. Jensen and L. L. Colgin, "Cross-frequency coupling between neuronal oscillations," *Trends in Cognitive Sciences*, vol. 11, no. 7, pp. 267 – 269, 2007.

[4] J. F. Hipp, D. J. Hawellek, M. Corbetta, M. Siegel, and A. K. Engel, "Large-scale cortical correlation structure of spontaneous oscillatory activity," *Nature Neuroscience*, vol. 15, no. 6, pp. 884–890, 2012.

[5] A. K. Engel, C. Gerloff, C. C. Hilgetag, and G. Nolte, "Intrinsic coupling modes: Multiscale interactions in ongoing brain activity," *Neuron*, vol. 80, no. 4, pp. 867–886, 2013.

[6] M. Siegel, T. H. Donner, and A. K. Engel, "Spectral fingerprints of large-scale neuronal interactions," *Nature Reviews Neuroscience*, vol. 13, no. 2, pp. 121–134, 2012.

[7] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, ser. Adaptive and Learning Systems for Signal Processing, Communications and Control Series. Wiley, 2001.

[8] S. Dähne, F. C. Meinecke, S. Haufe, J. Höhne, M. Tangermann, K.-R. Müller, and V. V. Nikulin, "SPoC: a novel framework for relating the amplitude of neuronal oscillations to behaviorally relevant parameters," *NeuroImage*, vol. 86, no. 0, pp. 111–122, 2014.

[9] S. Dähne, F. Bießman, F. C. Meinecke, J. Mehnert, S. Fazli, and K.-R. Müller, "Integration of multivariate data streams with bandpower signals," *IEEE Transactions on Multimedia*, vol. 15, no. 5, pp. 1001–1013, 2013.

[10] S. Baillet, J. C. J. Mosher, and R. M. R. Leahy, "Electromagnetic brain mapping," *IEEE Signal Processing Magazine*, vol. 18, no. 6, pp. 14–30, 2001.

[11] L. C. Parra, C. D. Spence, A. D. Gerson, and P. Sajda, "Recipes for the linear analysis of EEG." *NeuroImage*, vol. 28, no. 2, pp. 326–41, Nov. 2005.

[12] S. Haufe, F. Meinecke, K. Görgen, S. Dähne, J.-D. Haynes, B. Blankertz, and F. Bießmann, "On the interpretation of weight vectors of linear models in multivariate neuroimaging," *NeuroImage*, vol. 87, pp. 96–110, 2014.

[13] J. Nocedal, "Updating Quasi-Newton Matrices with Limited Storage," *Mathematics of Computation*, vol. 35, no. 151, pp. 773–782, 1980.

[14] D. Ramírez, P. J. Schreier, J. Vía, and V. V. Nikulin, "Power-CCA: Maximizing the correlation coefficient between the power of projections," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013.

[15] V. V. Nikulin, G. Nolte, and G. Curio, "Cross-frequency decomposition: A novel technique for studying interactions between neuronal oscillations with different frequencies," *Clinical Neurophysiology*, vol. 123, no. 7, pp. 1353–1360, 2012.

[16] G. Nolte, F. C. Meinecke, A. Ziehe, and K.-R. Müller, "Identifying interactions in mixed and noisy complex systems," *Phys Rev E*, vol. 73, p. 051913, 2006.

[17] M. Dyrholm, S. Makeig, and L. K. Hansen, "Model selection for convolutive ICA with an application to spatiotemporal analysis of EEG," *Neural Comput*, vol. 19, pp. 934–955, 2007.

[18] G. Gómez-Herrero, M. Atienza, K. Egiazarian, and J. L. Cantero, "Measuring directional coupling between EEG sources," *NeuroImage*, vol. 43, pp. 497–508, 2008.

[19] S. Haufe, R. Tomioka, G. Nolte, K.-R. Müller, and M. Kawanabe, "Modeling sparse connectivity between underlying brain sources for EEG/MEG," *IEEE Trans Biomed Eng*, vol. 57, pp. 1954–1963, 2010.

[20] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3/4, pp. 321–377, 1936.

[21] F. Bießmann, F. C. Meinecke, A. Gretton, A. Rauch, G. Rainer, N. K. Logothetis, and K.-R. Müller, "Temporal kernel cca and its application in multimodal neuronal data analysis," *Machine Learning*, vol. 79, no. 1-2, pp. 5–27, 2010.

[22] F. Bießmann, S. M. Plis, F. C. Meinecke, T. Eichele, and K.-R. Müller, "Analysis of multimodal neuroimaging data," *IEEE Reviews in Biomedical Engineering*, vol. 4, pp. 26–58, 2011.

[23] J. P. Dmochowski, P. Sajda, J. Dias, and L. C. Parra, "Correlated components of ongoing EEG point to emotionally laden attention - a possible marker of engagement?" *Front Hum Neurosci*, vol. 6, p. 112, 2012.