POWER-CCA: MAXIMIZING THE CORRELATION COEFFICIENT BETWEEN THE POWER OF PROJECTIONS

David Ramírez[1], *Peter J. Schreier*[1], *Javier Vía*[2] *and Vadim V. Nikulin*[3,4]

[1] Signal and System Theory Group, Universität Paderborn, Germany
e-mail: {david.ramirez,peter.schreier}@sst.upb.de
[2] Advanced Signal Processing Group, University of Cantabria, Santander, Spain
e-mail: jvia@gtas.dicom.unican.es
[3] Neurophysics Group, Charité – University Medicine, Berlin, Germany
[4] Bernstein Center for Computational Neuroscience, Berlin, Germany
e-mail: vadim.nikulin@charite.de

ABSTRACT

This work presents a variation of canonical correlation analysis (CCA), where the correlation coefficient between the instantaneous power of the projections is maximized, rather than between the projections themselves. The resulting optimization problem is not convex, and we have to resort to a sub-optimal approach. Concretely, we propose a two-step solution consisting of the singular value decomposition (SVD) of a "coherence" matrix followed by a rank-one matrix approximation. This technique is applied to blindly recovering signals in a model that is motivated by the study of neuronal dynamics in humans using electroencephalography (EEG) and magnetoencephalography (MEG). A distinctive feature of this model is that it allows recovery of amplitude-amplitude coupling between neuronal processes.

Index Terms— Bi-quadratic optimization, canonical correlation analysis (CCA), neuronal dynamics, electroencephalography (EEG), magnetoencephalography (MEG)

1. INTRODUCTION

In this paper, we develop a variation of canonical correlation analysis (CCA), called "Power-CCA." This technique is applied to blindly recovering, from two different linear mixtures, two sets of signals whose amplitudes are correlated. Such a setup is motivated by the study of neuronal dynamics in humans using electroencephalography (EEG) and magnetoencephalography (MEG). Conventionally, neuronal interactions are studied assuming phase synchronization between processes in the same frequency range [1, 2]. Recently, however, neuronal interactions have also been shown to occur at different frequency ranges (so-called cross-frequency interactions). These interactions include phase-phase, amplitude-phase and amplitude-amplitude coupling between the neuronal processes, which do not have the same frequency [3].

A majority of previous neuroimaging studies on cross-frequency interactions was based on an analysis in sensor space [3–6]. However, this approach has the disadvantage that source-mixing due to volume conduction can strongly obscure true topographic relationships between the interacting systems. Since EEG and MEG recordings are usually based on multichannel setups, an alternative approach is to find spatial filters that maximize the desired type of interactions. Previously such a decomposition has been introduced for studying phase-phase cross-frequency interactions [7]. In the present study, we introduce a novel multivariate decomposition technique for extracting neuronal components that exhibit amplitude-amplitude interactions between neuronal processes with different frequencies. To the best of our knowledge, this problem has never been considered in the literature.

Our technique is similar to canonical correlation analysis (CCA) [8]. The objective of CCA is to maximize the correlation coefficient between the projections (linear combinations) of two sets of variables. However, when looking for amplitude-amplitude correlation, CCA fails dramatically. We present Power-CCA, where the objective is to maximize the correlation coefficient between the instantaneous powers of the projections, rather than between the projections themselves. Since this leads to a non-convex optimization problem, we propose a sub-optimal solution. Its performance is illustrated using simulations that model oscillatory neuronal processes and amplitude-amplitude synchronization/correlation between them.

2. PROBLEM FORMULATION

The problem in neuroimaging that we are interested in can be described by the following simple model

$$\mathbf{x}[n] = \mathbf{A}\mathbf{u}[n] + \mathbf{n}_x[n]$$
$$\mathbf{y}[n] = \mathbf{B}\mathbf{v}[n] + \mathbf{n}_y[n],$$

where $\mathbf{A} \in \mathbb{R}^{L_x \times L_u}$ and $\mathbf{B} \in \mathbb{R}^{L_y \times L_v}$ are two fixed, but unknown, mixing matrices, assumed to have full rank, $\mathbf{u}[n] = [u_1[n], \ldots, u_{L_u}[n]]$ and $\mathbf{v}[n] = [v_1[n], \ldots, v_{L_v}[n]]$ are two sets of discrete-time signals, which contain the signals of interest, and $\mathbf{n}_x[n] \in \mathbb{R}^{L_x}$ and $\mathbf{n}_y[n] \in \mathbb{R}^{L_y}$ are zero-mean additive noises with unknown covariance matrices, uncorrelated with the signals. We assume $L_x \geq L_u$ and $L_y \geq L_v$.

The distinctive feature of this setup is that there is a strong linear coupling (i.e., high correlation) between some of the amplitudes of $\{u_i[n]\}_{i=1}^{L_u}$ and some of the amplitudes of $\{v_i[n]\}_{i=1}^{L_v}$. These amplitudes are defined as the absolute value of the corresponding analytic signals

$$u_i^+[n] = u_i[n] + j\mathcal{H}\{u_i[n]\}, \quad i = 1, \ldots, L_u,$$
$$v_i^+[n] = v_i[n] + j\mathcal{H}\{v_i[n]\}, \quad i = 1, \ldots, L_v,$$

where $\mathcal{H}\{\cdot\}$ denotes the discrete-time Hilbert transform. In order to measure the strength of the linear coupling between the amplitudes $|u_i^+[n]|$ and $|v_i^+[n]|$, we define the correlation coefficient

$$\rho_{|u_i^+|,|v_i^+|} = \frac{\sigma_{|u_i^+|,|v_i^+|}}{\sigma_{|u_i^+|}\sigma_{|v_i^+|}},$$

where the cross-covariance is

$$\sigma_{|u_i^+|,|v_i^+|} = E\left[\left|u_i^+[n]\right|\left|v_i^+[n]\right|\right] - E\left[\left|u_i^+[n]\right|\right]E\left[\left|v_i^+[n]\right|\right]$$

and the variances are

$$\sigma^2_{|u_i^+|} = E\left[\left|u_i^+[n]\right|^2\right] - E^2\left[\left|u_i^+[n]\right|\right]$$

and

$$\sigma^2_{|v_i^+|} = E\left[\left|v_i^+[n]\right|^2\right] - E^2\left[\left|v_i^+[n]\right|\right].$$

In our setup, we ignore any temporal correlation that the signals may or may not have. Thus, we will drop the time index for notational convenience. We further assume that the amplitudes $\{|u_i^+|\}$ are pairwise uncorrelated, the amplitudes $\{|v_i^+|\}$ are pairwise uncorrelated, and $|u_i^+|$ and $|v_j^+|$ are uncorrelated for $i \neq j$. However, the first $P \leq \min(L_u, L_v)$ amplitude pairs $|u_i^+|$ and $|v_i^+|$ are correlated, i.e., the correlation coefficient $\rho_{|u_i^+|,|v_i^+|}$ is nonzero for the first P pairs of $|u_i^+|$ and $|v_i^+|$. We would like to utilize this amplitude-amplitude correlation in order to recover the signals u_i and v_i, $i = 1, ..., P$, from the observed mixtures $\mathbf{x}$ and $\mathbf{y}$. Without loss of generality, we will assume that the signal pairs are ordered such that u_1 and v_1 are the most strongly correlated, u_2 and v_2 the second most strongly correlated, and so on. With this setup, one might think that any blind source separation technique would give us $\mathbf{u}[n]$ and $\mathbf{v}[n]$. However, this does not work for two reasons:

1. It would not tell us which signal $u_i[n]$ correlates with which signal $v_j[n]$.
2. It would fail if the signals are Gaussian.

3. POWER-CCA

Consider the analytic versions of the observed mixtures, which we will simply denote by $\mathbf{x}$ and $\mathbf{y}$, dropping the superscript $^+$ for notational convenience. In order to find the first pair of signals, we need to determine linear transformations $\mathbf{w}_x$ and $\mathbf{w}_y$ such that the correlation coefficient between the magnitudes of the projections $\xi_x = \mathbf{w}_x^H \mathbf{x}$ and $\xi_y = \mathbf{w}_y^H \mathbf{y}$ is maximized:

$$\underset{\mathbf{w}_x, \mathbf{w}_y}{\text{maximize}}\, \rho_{|\xi_x|,|\xi_y|}. \tag{1}$$

This problem is reminiscent of canonical correlation analysis (CCA), where we would solve

$$\underset{\mathbf{w}_x, \mathbf{w}_y}{\text{maximize}}\, \rho_{\xi_x,\xi_y}.$$

Nevertheless, solving (1) seems very difficult. As an easier alternative, we propose to maximize the correlation coefficient between the instantaneous power instead, that is,

$$\underset{\mathbf{w}_x, \mathbf{w}_y}{\text{maximize}}\, \rho_{|\xi_x|^2,|\xi_y|^2} = \underset{\mathbf{w}_x, \mathbf{w}_y}{\text{maximize}}\, \frac{\sigma_{|\xi_x|^2,|\xi_y|^2}}{\sigma_{|\xi_x|^2}\sigma_{|\xi_y|^2}},$$

where $\sigma_{|\xi_x|^2|\xi_y|^2}$ is the cross-covariance between the power of the projections, and $\sigma^2_{|\xi_x|^2}$ and $\sigma^2_{|\xi_y|^2}$ are the variances. We call this setup "Power-CCA."

Like in the original CCA problem, the directions may be found as the solution to the following optimization problem

$$\begin{aligned} \underset{\mathbf{w}_x, \mathbf{w}_y}{\text{maximize}}\, & \sigma_{|\xi_x|^2,|\xi_y|^2}, \\ \text{subject to } & \sigma^2_{|\xi_x|^2} = 1, \\ & \sigma^2_{|\xi_y|^2} = 1. \end{aligned} \tag{2}$$

The cross-covariance is given by

$$\begin{aligned} \sigma_{|\xi_x|^2,|\xi_y|^2} &= E\left[|\xi_x|^2|\xi_y|^2\right] - E\left[|\xi_x|^2\right]E\left[|\xi_y|^2\right] \\ &= E\left[|\xi_x|^2|\xi_y|^2\right] - \left(\mathbf{w}_x^H\mathbf{R}_{xx}\mathbf{w}_x\right)\left(\mathbf{w}_y^H\mathbf{R}_{yy}\mathbf{w}_y\right), \end{aligned}$$

where $\mathbf{R}_{xx}$ and $\mathbf{R}_{yy}$ are the cross-covariance matrices of $\mathbf{x}$ and $\mathbf{y}$, respectively, and the cross-correlation $E\left[|\xi_x|^2|\xi_y|^2\right]$ may be expressed as

$$\begin{aligned} E\left[|\xi_x|^2|\xi_y|^2\right] &= E\left[\left(\mathbf{w}_x^H\mathbf{x}\right)\left(\mathbf{w}_x^H\mathbf{x}\right)^*\left(\mathbf{w}_y^H\mathbf{y}\right)\left(\mathbf{w}_y^H\mathbf{y}\right)^*\right] \\ &= \bar{\mathbf{w}}_x^H\bar{\mathbf{R}}_{xy}\bar{\mathbf{w}}_y. \end{aligned}$$

In this equation, $\bar{\mathbf{R}}_{xy} = E[(\mathbf{x}\otimes\mathbf{x}^*)(\mathbf{y}\otimes\mathbf{y}^*)^H]$, $\bar{\mathbf{w}}_x = \mathbf{w}_x \otimes \mathbf{w}_x^*$, $\bar{\mathbf{w}}_y = \mathbf{w}_y \otimes \mathbf{w}_y^*$, where $\otimes$ denotes the Kronecker product. We may thus rewrite (2) as

$$\begin{aligned} \underset{\bar{\mathbf{w}}_x, \bar{\mathbf{w}}_y, \mathbf{w}_x, \mathbf{w}_y}{\text{maximize}}\, & \bar{\mathbf{w}}_x^H\bar{\mathbf{C}}_{xy}\bar{\mathbf{w}}_y, \\ \text{subject to } & \bar{\mathbf{w}}_x^H\bar{\mathbf{C}}_{xx}\bar{\mathbf{w}}_x = 1, \\ & \bar{\mathbf{w}}_y^H\bar{\mathbf{C}}_{yy}\bar{\mathbf{w}}_y = 1, \\ & \bar{\mathbf{w}}_x = \mathbf{w}_x \otimes \mathbf{w}_x^*, \\ & \bar{\mathbf{w}}_y = \mathbf{w}_y \otimes \mathbf{w}_y^*, \end{aligned} \tag{3}$$

where $\bar{\mathbf{C}}_{xy} = \bar{\mathbf{R}}_{xy} - \mathbf{r}_{xx}^{*}\mathbf{r}_{yy}^{T}$, $\bar{\mathbf{C}}_{xx} = \bar{\mathbf{R}}_{xx} - \mathbf{r}_{xx}^{*}\mathbf{r}_{xx}^{T}$, with $\bar{\mathbf{R}}_{xx} = E[(\mathbf{x}\otimes\mathbf{x}^{*})\,(\mathbf{x}\otimes\mathbf{x}^{*})^{H}]$ and $\mathbf{r}_{xx} = \text{vec}\,(\mathbf{R}_{xx})$. The terms $\bar{\mathbf{C}}_{yy}$, $\bar{\mathbf{R}}_{yy}$ and $\mathbf{r}_{yy}$ are defined analogously.

In the optimization problem (3), we maximize a bi-quadratic objective function subject to quartic constraints. Similar optimization problems have been previously considered by [9–12]. The problem (3) is not convex due to the Kronecker structure of $\bar{\mathbf{w}}_x$ and $\bar{\mathbf{w}}_y$. However, we may convexify it by dropping the Kronecker constraint, which then yields a sub-optimal solution. Doing so, the optimization problem is now equivalent to that of CCA [13], and its solution is given by the whitened singular vectors of the coherence matrix $\bar{\mathbf{Q}}_{xy} = \bar{\mathbf{C}}_{xx}^{-1/2}\bar{\mathbf{C}}_{xy}\bar{\mathbf{C}}_{yy}^{-1/2}$. That is, $\bar{\mathbf{w}}_x = \bar{\mathbf{C}}_{xx}^{-1/2}\mathbf{u}$ and $\bar{\mathbf{w}}_y = \bar{\mathbf{C}}_{yy}^{-1/2}\mathbf{v}$, where $\mathbf{u}$ and $\mathbf{v}$ are the left and right principal singular vectors of $\bar{\mathbf{Q}}_{xy}$.

Given the *unconstrained* solution, the Kronecker structure may be imposed by $\text{vec}\left(\mathbf{w}_x^{*}\mathbf{w}_x^{T}\right) = \mathbf{w}_x \otimes \mathbf{w}_x^{*}$. Defining the matrices $\bar{\mathbf{W}}_x = \text{unvec}(\bar{\mathbf{w}}_x)$ and $\bar{\mathbf{W}}_y = \text{unvec}(\bar{\mathbf{w}}_y)$, and their principal left singular vectors $\mathbf{s}_x$ and $\mathbf{s}_y$, the *approximate* solution to the problem is therefore

$$\mathbf{w}_x = \mathbf{s}_x^{*}, \qquad \mathbf{w}_y = \mathbf{s}_y^{*}.$$

If the matrices $\bar{\mathbf{W}}_x$ and $\bar{\mathbf{W}}_y$ were already rank-one, this solution would be optimal. By comparing the largest singular value of $\bar{\mathbf{W}}_x$ (and $\bar{\mathbf{W}}_y$) with the remaining singular values, we may therefore check how close the sub-optimal solution is to the optimal solution.

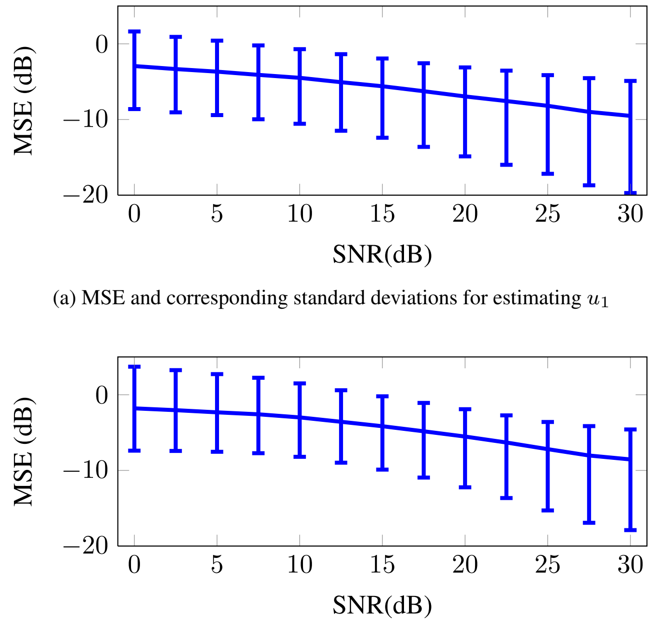
As with CCA, further projections can be obtained subject to the constraint that these be uncorrelated with previously obtained projections:

$$\sigma_{|\xi_x^{(i)}|^2|\xi_x^{(j)}|^2} = 0, \qquad \sigma_{|\xi_y^{(i)}|^2|\xi_y^{(j)}|^2} = 0,$$
$$\sigma_{|\xi_x^{(i)}|^2|\xi_y^{(j)}|^2} = 0, \qquad \sigma_{|\xi_y^{(i)}|^2|\xi_x^{(j)}|^2} = 0,$$

for $j = 1, \dots, i-1$. This results in the optimization problem

$$\begin{aligned}
\underset{\bar{\mathbf{w}}_x^{(i)},\bar{\mathbf{w}}_y^{(i)},\mathbf{w}_x^{(i)},\mathbf{w}_y^{(i)}}{\text{maximize}} \quad & \bar{\mathbf{w}}_x^{(i)H}\bar{\mathbf{C}}_{xy}\bar{\mathbf{w}}_y^{(i)}, && (4)\\
\text{subject to } & \bar{\mathbf{w}}_x^{(i)H}\bar{\mathbf{C}}_{xx}\bar{\mathbf{w}}_x^{(i)} = 1,\\
& \bar{\mathbf{w}}_y^{(i)H}\bar{\mathbf{C}}_{yy}\bar{\mathbf{w}}_y^{(i)} = 1,\\
& \bar{\mathbf{w}}_x^{(i)H}\bar{\mathbf{C}}_{xx}\bar{\mathbf{w}}_x^{(j)} = 0, \quad j = 1,\dots,i-1,\\
& \bar{\mathbf{w}}_x^{(i)H}\bar{\mathbf{C}}_{xy}\bar{\mathbf{w}}_y^{(j)} = 0, \quad j = 1,\dots,i-1,\\
& \bar{\mathbf{w}}_y^{(i)H}\bar{\mathbf{C}}_{yy}\bar{\mathbf{w}}_y^{(j)} = 0, \quad j = 1,\dots,i-1,\\
& \bar{\mathbf{w}}_y^{(i)H}\bar{\mathbf{C}}_{xy}^{H}\bar{\mathbf{w}}_x^{(j)} = 0, \quad j = 1,\dots,i-1,\\
& \bar{\mathbf{w}}_x^{(j)} = \mathbf{w}_x^{(j)} \otimes \mathbf{w}_x^{(j)*}, \quad j = 1,\dots,i,\\
& \bar{\mathbf{w}}_y^{(j)} = \mathbf{w}_y^{(j)} \otimes \mathbf{w}_y^{(j)*}, \quad j = 1,\dots,i.
\end{aligned}$$

Again, dropping the Kronecker structure constraint, the solution to the problem (4) is given by the ith left and right principal singular vectors of $\bar{\mathbf{Q}}_{xy}$, and the Kronecker structure may be imposed retrospectively using the same procedure as above.



(a) MSE and corresponding standard deviations for estimating u_1

(b) MSE and corresponding standard deviations for estimating u_2

Fig. 1: Results for length $N = 1000$

Given the projections, the analytic versions of the signals of interest may now be recovered as

$$\hat{u}_i^{+}[n] = [\mathbf{w}_x^{(i)}]^{H}\mathbf{x}[n], \qquad \hat{v}_i^{+}[n] = [\mathbf{w}_y^{(i)}]^{H}\mathbf{y}[n],$$

and the corresponding real signals are therefore

$$\hat{u}_i[n] = \text{Re}\{[\mathbf{w}_x^{(i)}]^{H}\mathbf{x}[n]\}, \qquad \hat{v}_i[n] = \text{Re}\{[\mathbf{w}_y^{(i)}]^{H}\mathbf{y}[n]\}.$$

Note that we can recover the analytic signals only up to an unknown complex number, which is typically the case in blind source separation.

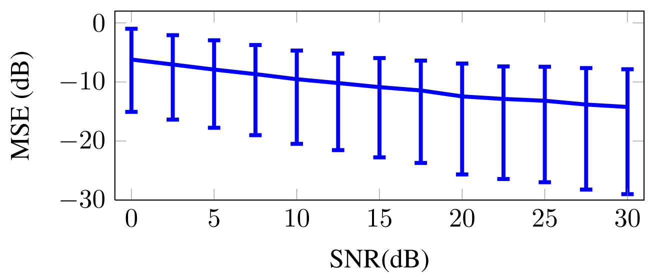
4. NUMERICAL RESULTS

In this section, we evaluate the performance of the proposed technique using Monte Carlo simulations. In particular, we obtain the mean square error (MSE) and its standard deviation. Since we cannot recover the phase and amplitude of the analytic signals, we define the mean square error as

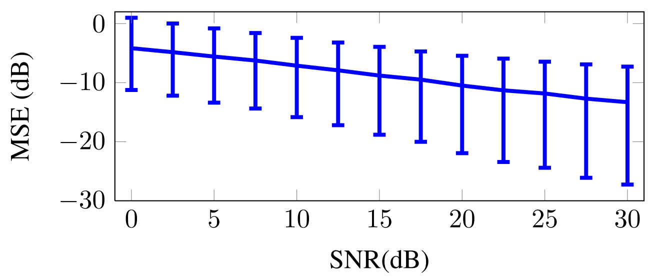
$$\text{MSE}_u\ (\text{dB}) = 10\log\left(\frac{1}{N}\sum_{n=1}^{N} E\left[\left|u_i[n] - \alpha\hat{u}_i[n]\right|^2\right]\right),$$

where α minimizes the error between $u_i^{+}[n]$ and $\hat{u}_i^{+}[n]$ and N denotes the length of the signals. The MSE of the second set is defined analogously by replacing u with v.

The first set of signals is made up of four independent bandpass Gaussian signals, centered at 10 Hz and sampled at



(a) MSE and standard deviations in the estimation of u_1



(b) MSE and standard deviations in the estimation of u_2

Fig. 2: Results for length $N = 10000$

100 Hz, and a fifth signal that is white Gaussian noise. The second set of signals consists of four independent bandpass Gaussian signals, centered at 25 Hz and sampled at 100 Hz, and a fifth signal that is white noise. So $L_u = L_v = 5$. The first two signal pairs in these two sets have an identical envelope, i.e., they have perfect amplitude-amplitude correlation, so $P = 2$. Both signal sets are mixed by square mixing matrices ($L_x = L_y = 5$), whose elements are independent and identically distributed and drawn from a Rayleigh distribution. Nothing is assumed known in these experiments, and the covariance matrices and other moments are estimated from the signals.

Figure 1 shows the MSE and corresponding standard deviation vs. the signal-to-noise ratio (SNR) for signal length $N = 1000$. We observe that the performance of recovering the first signal u_1 is slightly better than that of recovering u_2. This may be explained by the fact that the two projections are not completely uncorrelated due to dropping the Kronecker constraints. Figure 2 shows results for the same setup, but with a longer signal of length $N = 10000$. As expected, this improves performance because the sample covariance matrices are then a better estimate of the true covariance matrices. We should point out that the results of CCA are not shown because CCA completely fails to recover the signals. Finally, one comment is in order. In a typical EEG or MEG setup, the length of the recorded signals may even be up to 20 minutes. Sampled at 100 Hz, this yields up to $N = 120000$ samples, which significantly improves performance.

Finally, Figure 3 shows the distribution of the ratio between the largest eigenvalue and the trace of $\bar{\mathbf{W}}_x$ for the first solution, where the SNR is 30 dB, $N = 10000$ samples, and the number of realization is 10000. As can be seen, most of the times this ratio is almost one, which implies that our sub-optimal approach is close to the optimal solution.

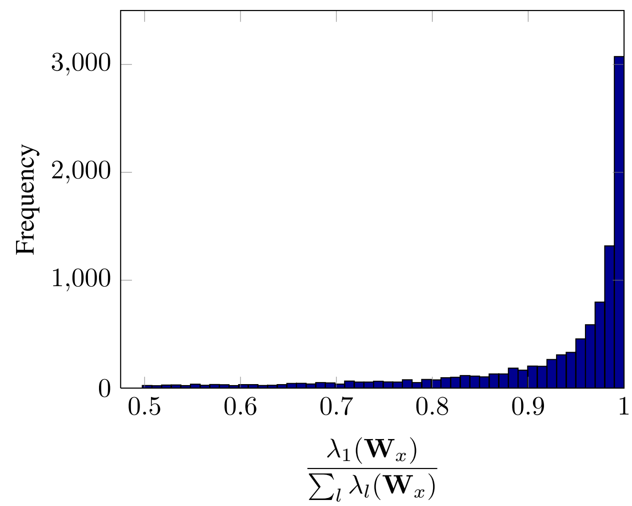


Fig. 3: Ratio between the largest singular value and the sum of all singular values for the first solution (SNR = 30 dB and $N = 10000$)

5. CONCLUSIONS

In this work, we have considered the problem of recovering, from two different linear mixtures, two sets of signals whose amplitudes are correlated. Our approach is closely related to canonical correlation analysis, but instead of maximizing the correlation coefficient between the projections, we maximize the correlation coefficient between their instantaneous powers. We have presented some simulations that model oscillatory neuronal processes and amplitude-amplitude synchronization/correlation between them. In this conference publication, we have focussed on the mathematical derivation. In a forthcoming journal publication, we will apply our technique to real EEG or MEG data.

6. ACKNOWLEDGMENTS

The work of P. J. Schreier was supported by the Alfried Krupp von Bohlen und Halbach Foundation, under its program “Return of German scientists from abroad”. The work of J. Vía was supported by the Spanish Government, Ministerio de Ciencia e Innovación (MICINN), under project COSIMA (TEC2010-19545-C04-03) and project COMONSENS (CSD2008-00010, CONSOLIDER-INGENIO 2010 Program). V. V. Nikulin acknowledges a partial support by the Bernstein Center for Computational Neuroscience, Berlin.

7. REFERENCES

[1] P. L. Nunez, R. Srinivasan, A. F. Westdorp, R. S. Wijesinghe, D. M. Tucker, R. B. Silberstein, and P. J. Cadusch, "EEG coherency: I: statistics, reference electrode, volume conduction, laplacians, cortical imaging, and interpretation at multiple scales," *Electroencephalogr. Clin. Neurophysiol.*, vol. 103, no. 5, pp. 499–515, 1997.

[2] G. Nolte, O. Baim, L. Wheaton, Z. Mari, S. Vorbach, and M. Hallet, "Identifying true brain interaction from EEG data using the imaginary part of coherency," *Clin. Neurophysiol.*, vol. 115, no. 10, pp. 2292–2307, 2004.

[3] R. T. Canolty, E. Edwards, S. S. Dalal, M. Soltani, S. S. Nagarajan, H. E. Kirsch, M. S. Berger, N. M. Barbaro, and R. T. Knight, "High gamma power is phase-locked to theta oscillations in human neocortex," *Science*, vol. 313, no. 5793, pp. 1626–1628, 2006.

[4] T. Schanze and R. Eckhorn, "Phase correlation among rhythms present at different frequencies: spectral methods, application to microelectrode recordings from visual cortex and functional implications," *Int. Journal of Psychophysiol.*, vol. 26, pp. 171–189, 1997.

[5] P. Tass, M. G. Rosenblum, J. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, and H.-J. Freund, "Detection of n:m phase locking from noisy data: Application to magnetoencephalography," *Physical Review Letters*, vol. 81, pp. 3291–3294, Oct. 1998.

[6] J. M. Palva, S. Palva, and K. Kaila, "Phase synchrony among neuronal oscillations in the human cortex," *Journal of Neuroscience*, vol. 25, pp. 3962–3672, 2005.

[7] 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.

[8] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321–377, 1936.

[9] C. Ling, J. Nie, L. Qi, and Y. Ye, "Bi-quadratic optimization over unit spheres and semidefinite programming relaxations," *SIAM Journal on Optimization*, vol. 20, no. 3, pp. 1286–1310, 2010.

[10] I. M. Bomze, C. Ling, L. Qi, and X. Zhang, "Standard bi-quadratic optimization problems and unconstrained polynomial reformulations," *Journal of Global Optimization*, vol. 52, no. 4, pp. 663–687, Apr. 2012.

[11] Y. Wang, L. Qi, and X. Zhang, "A practical method for computing the largest Meigenvalue of a fourthorder partially symmetric tensor," *Numerical Linear Algebra with Applications*, vol. 16, no. 7, pp. 589–601, 2009.

[12] Z.-Q. Luo and S. Zhang, "A semidefinite relaxation scheme for multivariate quartic polynomial optimization with quadratic constraints," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1716–1736, Jan. 2010.

[13] Magnus Borga, *Learning Multidimensional Signal Processing*, Ph.D. thesis, Linköping University, Linköping, Sweden, 1998, Dissertation No 531.