KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Geirnaert S., Francart T., Bertrand A. (2023), **Stimulus-Informed Generalized Canonical Correlation Analysis of Stimulus-Following Brain Responses** *Proceedings of the 31st European Signal Processing Conference (EUSIPCO 2023)*
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	
Journal homepage	https://eusipco2023.org/
Author contact	simon.geirnaert@esat.kuleuven.be *+ 32 (0)16 37 35 36*
Abstract	
IR	

(article begins on next page)

Stimulus-Informed Generalized Canonical Correlation Analysis of Stimulus-Following Brain Responses

Simon Geirnaert[1,2,3], Tom Francart[2,3], and Alexander Bertrand[1,3]

[1]*KU Leuven, Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics*, Leuven, Belgium
[2]*KU Leuven, Department of Neurosciences, ExpORL*, Leuven, Belgium
[3]*KU Leuven Institute for Artificial Intelligence (Leuven.AI)*, Leuven, Belgium
{simon.geirnaert, tom.francart, alexander.bertrand}@kuleuven.be

***Abstract*—In brain-computer interface or neuroscience applications, generalized canonical correlation analysis (GCCA) is often used to extract correlated signal components in the neural activity of different subjects attending to the same stimulus. This allows quantifying the so-called inter-subject correlation or boosting the signal-to-noise ratio of the stimulus-following brain responses with respect to other (non-)neural activity. GCCA is, however, stimulus-unaware: it does not take the stimulus information into account and does therefore not cope well with lower amounts of data or smaller groups of subjects. We propose a novel stimulus-informed GCCA algorithm based on the MAXVAR-GCCA framework. We show the superiority of the proposed stimulus-informed GCCA method based on the inter-subject correlation between electroencephalography responses of a group of subjects listening to the same speech stimulus, especially for lower amounts of data or smaller groups of subjects.**

***Index Terms*—generalized canonical correlation analysis, electroencephalography, stimulus-following neural response**

I. Introduction

Traditionally, several neuroscience applications and brain-computer interface technologies heavily rely on active participation of the user following specific instructions. Moreover, synthetic and controlled stimuli such as flickering visual patterns or beep tones are used, as they evoke more controlled brain responses. Lastly, they rely on multi-trial experiments, where the same controlled stimulus is repeatedly presented to allow enhancing the signal-to-noise ratio (SNR) by averaging the responses across multiple trials [1]. In the past few years, a shift has been occurring towards passive, single-trial experiments using natural sensory stimuli, such as speech or video footage [2]–[8]. This shift opens doors to novel applications, for example, in attention tracking in hearing devices [2], [3], the (online) classroom [4], neuromarketing [5], [6], or virtual reality environments [9]. Given that we are interested in stimulus-following brain responses, we focus on electroencephalography (EEG) or magnetoencephalography (MEG), which have an excellent temporal resolution [1].

This shift brings along several signal processing-related challenges, such as the strong subject-specificity arising from using uncontrolled and natural stimuli that evoke highly variable responses, and especially the low SNR of the stimulus-following EEG responses buried under interfering non-neural and neural activity that is not time-locked to the stimulus. In this single-trial setting, the low SNR can not be dealt with anymore by averaging multiple trials, such that data-driven filtering methods are required to enhance the stimulus-following neural response and suppress all other noise sources.

In this paper, we focus on generalized canonical correlation analysis (GCCA), which allows extracting correlated components across multiple (EEG) recordings, for example from different subjects. GCCA can then be used not only to enhance the SNR but also to quantify inter-subject correlation (ISC) [4]–[7], for dimensionality reduction, or to summarize a set of EEG recordings [10]. An overview of GCCA for brain data analysis can be found in de Cheveigné et al. [10].

We focus on the problem of extracting the stimulus-following neural response from a set of EEG recordings from multiple subjects attending to the same stimulus. One of the strengths of GCCA here is that it is *stimulus-unaware*: it makes no assumptions about the stimulus (representation) and also works when the stimulus is unavailable. However, this strength can be a weakness at the same time. As explained before, one of the main challenges is the notoriously low SNR of the stimulus-following neural response. Maximally exploiting all available side information, including the stimulus, is paramount to optimally boost the SNR, even more so when only a few subjects or small amount of data are available. The latter regularly occurs, for example, in a time-adaptive context for online processing [3]. Therefore, we modify the GCCA formulation to include the stimulus as side-information, leading to *stimulus-informed* GCCA (SI-GCCA).

This research is funded by a PDM mandate from KU Leuven (for S. Geirnaert, No PDMT1/22/009), FWO project nr. G081722N, the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 802895), and the Flemish Government (AI Research Program). The scientific responsibility is assumed by its authors.

In Section II, we revisit GCCA and, more specifically, its MAXVAR formulation. We then present the novel SI-GCCA method in Section III. In Section IV, we compare GCCA and SI-GCCA. Conclusions are drawn in Section V.

II. MAXVAR-GCCA

Several ways exist to generalize traditional CCA to multiple views or subjects, of which SUMCORR and MAXVAR are the most popular formulations [11]. Here we choose the MAXVAR-GCCA formulation because it allows for an easy introduction of the stimulus information (Section III) and has the nice property of leading to a generalized eigenvalue decomposition (GEVD) [11].

Consider K stimulus-driven zero-mean EEG signals $\mathbf{X}_k \in \mathbb{R}^{T\times M}$ of K subjects attending to the same stimulus, for example, a speech signal. T denotes the number of EEG samples, where each sample is M-dimensional[1], arising from, for example, different EEG channels and/or time-lagged copies of each channel. In MAXVAR-GCCA, the goal is to find the $M\times Q$-dimensional filters $\mathbf{W}_k \in \mathbb{R}^{M\times Q}$ and a Q-dimensional shared signal subspace spanned by $\mathbf{S} = \begin{bmatrix}\mathbf{s}_1 & \cdots & \mathbf{s}_Q\end{bmatrix} \in \mathbb{R}^{T\times Q}$ such that the latter is on average closest to the transformed individual EEG signals $\mathbf{X}_k\mathbf{W}_k$ [11]–[13]:

$$\begin{aligned}\min_{\mathbf{W}_1,\ldots,\mathbf{W}_K,\mathbf{S}} \quad & \sum_{k=1}^{K} ||\mathbf{S}-\mathbf{X}_k\mathbf{W}_k||_F^2 \\ \text{s.t.} \quad & \mathbf{S}^\mathsf{T}\mathbf{S} = \mathbf{I}_Q,\end{aligned} \tag{1}$$

with $\mathbf{I}_Q$ the Q-dimensional identity matrix.

The Lagrangian function is then equal to:

$$\begin{aligned}&\mathcal{L}(\mathbf{W}_1,\ldots,\mathbf{W}_K,\mathbf{S},\mathbf{\Lambda}) = \\ &\quad K\text{Tr}\left(\mathbf{S}^\mathsf{T}\mathbf{S}\right) - 2\sum_{k=1}^{K}\text{Tr}\left(\mathbf{S}^\mathsf{T}\mathbf{X}_k\mathbf{W}_k\right) + \\ &\quad \sum_{k=1}^{K}\text{Tr}\left(\mathbf{W}_k^\mathsf{T}\mathbf{X}_k^\mathsf{T}\mathbf{X}_k\mathbf{W}_k\right) - \text{Tr}\left(\left(\mathbf{S}^\mathsf{T}\mathbf{S}-\mathbf{I}_Q\right)\mathbf{\Lambda}\right),\end{aligned}$$

with $\mathbf{\Lambda}\in\mathbb{R}^{Q\times Q}$ a symmetric matrix containing the Lagrange multipliers. From the Karush-Kuhn-Tucker conditions, we find the following three equations:

$$\begin{cases}\mathbf{X}_k^\mathsf{T}\mathbf{S} = \mathbf{X}_k^\mathsf{T}\mathbf{X}_k\mathbf{W}_k, \ k=1,\ldots,K & (2)\\ \mathbf{S} = \sum_{k=1}^{K}\mathbf{X}_k\mathbf{W}_k\mathbf{\Omega}, & (3)\\ \mathbf{S}^\mathsf{T}\mathbf{S} = \mathbf{I}_Q, & (4)\end{cases}$$

with $\mathbf{\Omega} = \left(K\mathbf{I}_Q - \mathbf{\Lambda}\right)^{-1}$ also a symmetric matrix. Defining $\mathbf{R}_{kl} = \mathbf{X}_k^\mathsf{T}\mathbf{X}_l \in \mathbb{R}^{M\times M}$ as the sample crosscorrelation matrix of $\mathbf{X}_k$ and $\mathbf{X}_l$, and plugging (3) into (2), one obtains

$$\mathbf{R}_{D_{xx}}\mathbf{W} = \mathbf{R}_{xx}\mathbf{W}\mathbf{\Omega}, \tag{5}$$

with $\mathbf{W} = \begin{bmatrix}\mathbf{W}_1^\mathsf{T} & \cdots & \mathbf{W}_K^\mathsf{T}\end{bmatrix}^\mathsf{T} \in \mathbb{R}^{KM\times Q}$, $\mathbf{R}_{D_{xx}} = \text{Blkdiag}(\mathbf{R}_{11},\ldots,\mathbf{R}_{KK}) \in \mathbb{R}^{KM\times KM}$ the block diagonal matrix containing the per-subject autocorrelation matrices, and $\mathbf{R}_{xx} = \mathbf{X}^\mathsf{T}\mathbf{X} \in \mathbb{R}^{KM\times KM}$, with $\mathbf{X} = \begin{bmatrix}\mathbf{X}_1 & \cdots & \mathbf{X}_K\end{bmatrix} \in \mathbb{R}^{T\times KM}$.

Given that $\mathbf{\Omega}$ is a symmetric matrix, it is orthogonally diagonalizable as

$$\mathbf{\Omega} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^\mathsf{T}, \tag{6}$$

with $\mathbf{U}\in\mathbb{R}^{Q\times Q}$ an orthogonal matrix and $\mathbf{\Sigma}\in\mathbb{R}^{Q\times Q}$ a diagonal matrix. Substituting (6) in (5) leads to the GEVD that gives the solution:

$$\mathbf{R}_{D_{xx}}\mathbf{W}\mathbf{U} = \mathbf{R}_{xx}\mathbf{W}\mathbf{U}\mathbf{\Sigma},$$

The optimal filters in $\mathbf{W}$ are then in the subspace spanned by the Q generalized eigenvectors (GEVcs) corresponding to the smallest generalized eigenvalues (GEVls)[2]. As it can be shown that the solution is defined upon any orthogonal transformation [14][3], for convenience, we take the GEVcs themselves as the solution. Note that this corresponds to selecting the matrix of Lagrange multipliers $\mathbf{\Lambda}$, or $\mathbf{\Omega}$, to be diagonal. The correct scaling of the GEVcs can be retrieved from (4). Lastly, as shown in [12], [13], using (3), the problem in (1) is equivalent to minimizing the sum of the pairwise distances between $\mathbf{X}_k\mathbf{w}_k$ and $\mathbf{X}_l\mathbf{w}_l$ for all k and l, relating to the intuitive definition of GCCA, i.e., maximizing pairwise correlations.

III. SI-GCCA

In stimulus-informed GCCA (SI-GCCA), we now want to use the stimulus, if available, as side information to steer the estimation of the shared signal subspace and filters. Such a framework is based on the (reasonable) assumption that the correlated components across the K EEG recordings correspond to the neural responses to the common stimulus. Using the stimulus to inform the GCCA estimation can be viewed as a task-informed regularization technique to cope with situations where only little data is available (such as in an adaptive or online context) or only a few subjects are available.

Assume a P-dimensional stimulus representation $\mathbf{Y}\in\mathbb{R}^{T\times P}$ (see Section IV-C2 for an example of such a representation for a speech stimulus). We propose to introduce the stimulus information into the MAXVAR-GCCA optimization problem in (1) by adding a forward model/encoder $\mathbf{V}\in\mathbb{R}^{P\times Q}$ on the stimulus to the shared signal subspace as an extra term:

$$\begin{aligned}\min_{\mathbf{W}_1,\ldots,\mathbf{W}_K,\mathbf{V},\mathbf{S}} \quad & \sum_{k=1}^{K}||\mathbf{S}-\mathbf{X}_k\mathbf{W}_k||_F^2 + \rho\,||\mathbf{S}-\mathbf{Y}\mathbf{V}||_F^2 \\ \text{s.t.} \quad & \mathbf{S}^\mathsf{T}\mathbf{S} = \mathbf{I}_Q.\end{aligned} \tag{7}$$

This formulation has the advantage that it remains a MAXVAR-GCCA problem with the stimulus as an additional view, so it can still be solved as a GEVD. The additional term related to the stimulus 'pulls' on the shared signal subspace, where the hyperparameter ρ determines the weight.

The new MAXVAR-GCCA optimization problem in (7) can be solved similarly to Section II by defining the new variables

[1] In principle, a different dimensionality per subject can be used.

[2] It can be found that the objective function at the optimal solution is equal to $K - \text{Tr}\left(\mathbf{\Sigma}^{-1}\right)$.

[3] Proof omitted due to space constraints.

$\tilde{\mathbf{W}} = \begin{bmatrix}\mathbf{W}_1^{\mathrm{T}} & \cdots & \mathbf{W}_K^{\mathrm{T}} & \mathbf{V}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{(KM+P)\times Q}$ and $\tilde{\mathbf{X}} = \begin{bmatrix}\mathbf{X}_1 & \cdots & \mathbf{X}_K & \mathbf{Y}\end{bmatrix} \in \mathbb{R}^{T\times(KM+P)}$. We now a priori select the matrix of Lagrange multipliers $\mathbf{\Lambda}$ to be diagonal. The resulting shared signal subspace $\mathbf{S}$ and corresponding GEVD then become:

$$\begin{cases} \mathbf{S} = \left(\sum_{k=1}^{K} \mathbf{X}_k \mathbf{W}_k + \rho \mathbf{Y}\mathbf{V}\right) \mathbf{\Omega} & (8) \\ \mathbf{R}_{D_{\tilde{x}\tilde{x}}} \tilde{\mathbf{W}} = \tilde{\mathbf{R}}_{\tilde{x}\tilde{x}} \tilde{\mathbf{W}} \mathbf{\Omega}, & (9) \end{cases}$$

with $\mathbf{R}_{D_{\tilde{x}\tilde{x}}} = \mathrm{Blkdiag}(\mathbf{R}_{11}, \ldots, \mathbf{R}_{KK}, \mathbf{R}_{yy})$, $\mathbf{\Omega} = ((K+\rho)\mathbf{I}_Q - \mathbf{\Lambda})^{-1}$ again a diagonal matrix, and

$$\tilde{\mathbf{R}}_{\tilde{x}\tilde{x}} = \begin{bmatrix} \mathbf{R}_{11} & \ldots & \mathbf{R}_{1K} & \rho\mathbf{R}_{1y} \\ \vdots & & \vdots & \vdots \\ \mathbf{R}_{1K} & \ldots & \mathbf{R}_{KK} & \rho\mathbf{R}_{Ky} \\ \mathbf{R}_{1y} & \ldots & \mathbf{R}_{Ky} & \rho\mathbf{R}_{yy} \end{bmatrix}. \quad (10)$$

By taking the GEVcs corresponding to the smallest GEVls, the correct filter subspace can again be retrieved. The correct scaling of the GEVcs can, as before, be retrieved from the equality constraint $\mathbf{S}^{\mathrm{T}}\mathbf{S} = \mathbf{I}_Q$. The SI-GCCA algorithm is summarized in Algorithm 1, while a MATLAB implementation is available at https://github.com/AlexanderBertrandLab/si-gcca.

Algorithm 1: SI-GCCA

Input: K stimulus-driven EEG signals $\mathbf{X}_k \in \mathbb{R}^{T\times M}$, stimulus $\mathbf{Y} \in \mathbb{R}^{T\times P}$, hyperparameter ρ, subspace dimension Q
Output: Per-subject filters $\mathbf{W}_k \in \mathbb{R}^{M\times Q}$

1: Compute correlation matrices $\tilde{\mathbf{R}}_{\tilde{x}\tilde{x}}$ as in (10) and $\mathbf{R}_{D_{\tilde{x}\tilde{x}}} = \mathrm{Blkdiag}(\mathbf{R}_{11}, \ldots, \mathbf{R}_{KK}, \mathbf{R}_{yy})$, with $\mathbf{R}_{kl} = \mathbf{X}_k^{\mathrm{T}}\mathbf{X}_l$ (optionally with regularization)
2: Compute $\tilde{\mathbf{W}}$ as the Q GEVcs corresponding to the Q smallest GEVls of the matrix pencil $\left(\mathbf{R}_{D_{\tilde{x}\tilde{x}}}, \tilde{\mathbf{R}}_{\tilde{x}\tilde{x}}\right)$
3: Scale the GEVcs such that $\mathbf{S}^{\mathrm{T}}\mathbf{S} = \mathbf{I}_Q$, with $\mathbf{S}$ defined in (8)
4: Extract $\mathbf{W}_k$ from $\tilde{\mathbf{W}} = \begin{bmatrix}\mathbf{W}_1^{\mathrm{T}} & \cdots & \mathbf{W}_K^{\mathrm{T}} & \mathbf{V}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$

IV. Experiments and Results

A. Task

We evaluate and compare the GCCA and SI-GCCA methods in the context of a group of subjects listening to the same speech stimulus, i.e., a story. For simplicity, we compute only one filter per subject, i.e., $Q = 1$. In this case, the set of Q filters $\mathbf{W}_k$ per subject boil down to a single M-dimensional filter $\mathbf{w}_k \in \mathbb{R}^M$ per subject. The goal is to use (SI-)GCCA to quantify the inter-subject correlation (ISC), i.e., the correlation between the (transformed) EEG recordings of the different subjects, which could serve as a proxy for attentional engagement of the group [4]–[7]. Given the zero-mean (one-dimensional) (SI-)GCCA transformed EEG signals $\mathbf{z}_k = \mathbf{X}_k\mathbf{w}_k \in \mathbb{R}^T$, this ISC is defined as the average pairwise Pearson correlation coefficient:

$$\mathrm{ISC} = \frac{2}{K(K-1)} \sum_{k=1}^{K-1} \sum_{l=k+1}^{K} \frac{\mathbf{z}_k^{\mathrm{T}}\mathbf{z}_l}{||\mathbf{z}_k||_2 \, ||\mathbf{z}_l||_2}. \quad (11)$$

This ISC, evaluated on trials of 60 s on an independent test set, serves as the performance metric. The higher the ISC the better, as this means that the neural processes unrelated to the stimulus are more suppressed.

In Section IV-B, we briefly introduce the dataset, while the design choices are shown in Section IV-C. The GCCA and SI-GCCA method are compared in two different experiments: varying the amount of training data (Section IV-D) and the number of subjects (Section IV-E). MATLAB code for these experiments is available at https://github.com/AlexanderBertrandLab/si-gcca.

B. Dataset

The dataset is taken from Broderick et al. [15] and is publicly available [16]. We use the EEG data of the first experiment, containing data from 19 normal-hearing subjects, all listening to the same narrative speech, i.e., an audiobook. This audiobook was presented in 20 trials, each of length around 3 min. Each of these trials was cut into new trials of 60 s, resulting in 52 min of synchronized EEG/speech data per subject. The EEG is recorded using a 128-channel BioSemi ActiveTwo system and re-referenced to the average mastoid channel.

C. Design choices

1) Stimulus representation and bandpass filtering: Given that it has been repeatedly established that EEG signals track the low-frequency envelope of the attended speech signal [2], [3], [17], the speech stimulus in the SI-GCCA framework is represented by the amplitude envelope, computed using the Hilbert transform [17]. Furthermore, both the speech envelope and EEG signals are filtered in the δ-band, i.e., between 1–4 Hz, which has shown to provide good results in the context of tracking auditory attention to a speech signal [8]. The EEG signals and speech envelope are correspondingly downsampled to 8 Hz.

2) Filter design: We will estimate spatio-temporal filters on the $C = 128$-channel EEG signals of the different subjects, augmented with L time lags to compensate for temporal differences in processing between subjects. Therefore, $M = CL$, with $L = 5$, i.e., the time lags are ranging from $-\frac{L-1}{2} = -2$ to $\frac{L-1}{2} = 2$, corresponding to an integration window of $[-250, 250]$ms. The resulting EEG regression matrix $\mathbf{X}_k$ then becomes a block Hankel matrix:

$$\mathbf{X}_k = \begin{bmatrix}\mathbf{X}_{k,1} & \ldots & \mathbf{X}_{k,128}\end{bmatrix},$$

$$\mathbf{X}_{k,c} = \begin{bmatrix} 0 & 0 & x_{k,c}(0) & x_{k,c}(1) & x_{k,c}(2) \\ 0 & x_{k,c}(0) & x_{k,c}(1) & x_{k,c}(2) & x_{k,c}(3) \\ x_{k,c}(0) & x_{k,c}(1) & x_{k,c}(2) & x_{k,c}(3) & x_{k,c}(4) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{k,c}(T-3) & x_{k,c}(T-2) & x_{k,c}(T-1) & 0 & 0 \end{bmatrix},$$

with $x_{k,c}(t)$ the cth-channel EEG signal of the kth subject.

In the SI-GCCA estimation, the stimulus representation is the speech envelope and 10 time-lagged copies of it to allow for spectral filtering (including compensation for the intrinsic delay between the stimulus and EEG response) by $\mathbf{v}$, leading to a pre-stimulus integration window of $[-1.25, 0]$s [2]. Therefore, the stimulus representation $\mathbf{Y}$ is also a Hankel matrix with $P = 11$ columns.

3) Correlation matrix estimation: All 60 s-trials are a priori normalized by setting the mean per channel to zero and putting the Frobenius norm across all channels to one. Furthermore, all correlation matrices ($\mathbf{R}_{xx}, \tilde{\mathbf{R}}_{\tilde{x}\tilde{x}}, \dots$) are estimated using ridge regression, where the regularization parameter is automatically estimated using the method proposed by Ledoit and Wolf [18].

4) Hyperparameter selection: The hyperparameter ρ in the SI-GCCA estimation problem (7), which determines the weight put on the stimulus, is selected based on the ISC using an independent validation set consisting of 20% of the trials in the test set (which are thereafter removed from this test set and not used in the ISC computation). A sweep of hyperparameters in the range of $\{0, 10^{-1}, 10^{-0.5}, \dots, 10^{3}\}$ is performed and the ρ leading to the highest ISC on the validation set is selected.

5) Significance level computation and statistical comparison: In Section IV-D and IV-E, per ISC, the significance level is determined via a random permutation test: for the GCCA algorithm, a random 5-fold cross-validation is performed. The resulting GCCA output signals are then 10 000 times randomly permuted per 60 s-trial across the different subjects (i.e., removing correlated patterns), after which the 5% significance level can be computed. This significance level is mainly determined by the trial length and number of subjects used to compute the ISC.

The statistical comparison between the (SI-)GCCA methods is performed using a linear mixed-effect model (LMEM), where the method is modelled as a fixed effect and the amount of training data (Section IV-D) or number of subjects (Section IV-E) as random slope (i.e., the ISC can vary differently per method across the random effect):

$$\text{ISC} \sim 1 + \text{method} + \left(\text{method} \,\middle|\, \begin{matrix}\text{amount of training data} \\ \text{number of subjects}\end{matrix}\right).$$

D. Amount of training data

In this experiment, we compare the GCCA and SI-GCCA methods for various amounts of training data while using all 19 subjects to train and test. The amount of training data is varied between 1 and 45 min. Per amount of training data, 50 Monte Carlo runs are performed, in which the 60 s-training trials are randomly sampled from the 52 min available data. The remaining 60 s-trials are used to test the resulting decoders. In case of the SI-GCCA method, this rest is split into 80% test data and 20% validation data.

Figure 1 shows the ISC as a function of the amount of training data used. The SI-GCCA method outperforms the GCCA method across all amounts of training data (mean difference in ISC of 0.011), which is confirmed using the LMEM (significant difference, p-value < 0.001). This shows that the stimulus information can effectively improve the per-subject filter estimation. The difference becomes smaller only when large amounts of training data are used: the side-information introduced by the stimulus regularizer decreases as the high dimensionality of the problem can be compensated by the large amounts of training data.

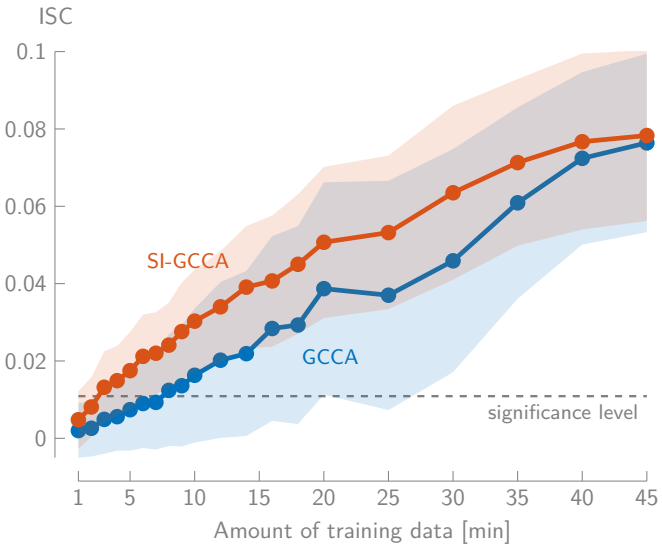

Fig. 1: SI-GCCA outperforms GCCA across all amounts of training data (mean $\pm$ standard deviation across all test trials and Monte Carlo runs).

Lastly, the SI-GCCA method allows retrieving significant ISCs already with 3 min of training data (as opposed to 8 min for the GCCA method), which is a crucial advantage in online, adaptive BCI applications.

E. Number of subjects

We also compare the GCCA and SI-GCCA methods for varying numbers of subjects while using 5-fold cross-validation on all available data (corresponding to 42 min of training data). Per number of subjects, 25 different combinations are sampled from all 19 subjects.

Figure 2 shows the ISC as a function of the number of subjects. The more subjects are used, the lower the significance level, as averaging more pairwise correlations in the ISC counteracts spurious correlations. The SI-GCCA method effectively compensates for the lower number of subjects, where less information is available to extract the shared stimulus-related activity. Using an LMEM, there is a significant difference between both methods across all numbers of subjects (mean difference in ISC of 0.023, p-value < 0.001). Furthermore, SI-GCCA results in significant correlations when more than 5 subjects are present, which is only the case for 7 subjects for the GCCA method.

V. Conclusion

In this paper, we have proposed a novel stimulus-informed GCCA framework that retains the attractive properties of the

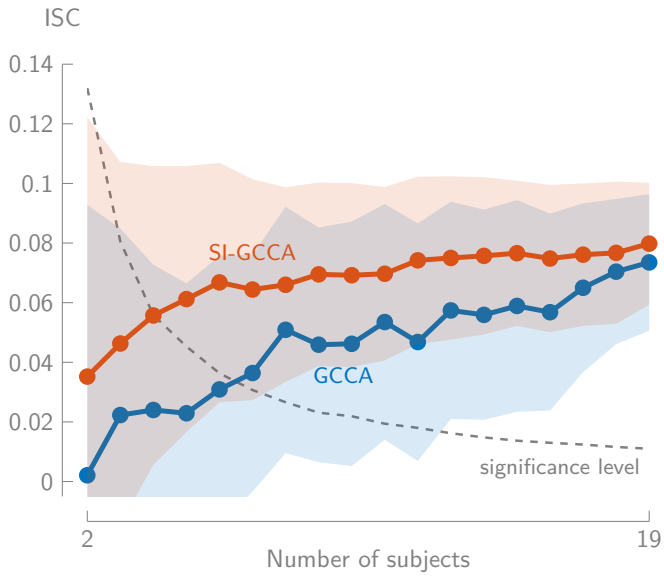

Fig. 2: SI-GCCA outperforms GCCA across all numbers of subjects (mean $\pm$ standard deviation across all folds and different sampled combinations).

GCCA-MAXVAR framework, i.e., it can still be computed as a GEVD. We have demonstrated its superiority w.r.t. GCCA when quantifying the ISC between subjects listening to the speech stimulus, especially when smaller amounts of training data and fewer subjects are available.

VI. Acknowledgements

The authors would like to thank Thomas Strypsteen and Miguel Bhagubai for their help in providing the necessary computing power and Charles Hovine for the brainstorming sessions and proofs about the invariance of the solution upon an orthogonal transformation.

References

[1] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain Computer Interfaces, a Review," *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2012.

[2] S. Geirnaert, S. Vandecappelle, E. Alickovic, A. de Cheveigné, E. C. Lalor, B. T. Meyer, S. Miran, T. Francart, and A. Bertrand, "Electroencephalography-Based Auditory Attention Decoding: Toward Neurosteered Hearing Devices," *IEEE Signal Processing Magazine*, vol. 38, no. 4, pp. 89–102, 2021.

[3] S. Geirnaert, T. Francart, and A. Bertrand, "Time-adaptive Unsupervised Auditory Attention Decoding Using EEG-based Stimulus Reconstruction," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 8, pp. 3767–3778, 2022.

[4] A. T. Poulsen, S. Kamronn, J. P. Dmochowski, L. C. Parra, and L. K. Hansen, "EEG in the classroom: Synchronised neural recordings during video presentation," *Scientific Reports*, vol. 7, pp. 1–9, 2017.

[5] 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?" *Frontiers in Human Neuroscience*, vol. 6, p. 112, 2012.

[6] J. P. Dmochowski, M. A. Bezdek, B. P. Abelson, J. S. Johnson, E. H. Schumacher, and L. C. Parra, "Audience preferences are predicted by temporal reliability of neural processing," *Nature Communications*, vol. 5, no. 1, pp. 1–9, 2014.

[7] J. J. Ki, S. P. Kelly, and L. C. Parra, "Attention strongly modulates reliability of neural responses to naturalistic narrative stimuli," *Journal of Neuroscience*, vol. 36, no. 10, pp. 3092–3101, 2016.

[8] J. Vanthornhout, L. Decruy, and T. Francart, "Effect of Task and Attention on Neural Tracking of Speech," *Frontiers in Neuroscience*, vol. 13, p. 977, 2019.

[9] V. Delvigne, H. Wannous, J.-P. Vandeborre, L. Ris, and T. Dutoit, "Attention Estimation in Virtual Reality with EEG based Image Regression," in *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, 2020, pp. 10–16.

[10] A. de Cheveigné, G. M. Di Liberto, D. Arzounian, D. D. Wong, J. Hjortkjær, S. Fuglsang, and L. C. Parra, "Multiway canonical correlation analysis of brain data," *NeuroImage*, vol. 186, no. October 2018, pp. 728–740, 2019.

[11] J. R. Kettenring, "Canonical analysis of several sets of variables," *Biometrika*, vol. 58, no. 3, pp. 433–451, 1971.

[12] J. Vía, I. Santamaría, and J. Pérez, "A learning algorithm for adaptive canonical correlation analysis of several data sets," *Neural Networks*, vol. 20, no. 1, pp. 139–152, 2007.

[13] C. Hovine and A. Bertrand, "Distributed MAXVAR: Identifying Common Signal Components across the Nodes of a Sensor Network," in *2021 29th European Signal Processing Conference (EUSIPCO)*. EURASIP, 2021, pp. 2159–2163.

[14] B. Ghojogh, F. Karray, and M. Crowley, "Eigenvalue and Generalized Eigenvalue Problems: Tutorial," *arXiv*, no. 1903.11240, 2022.

[15] M. P. Broderick, A. J. Anderson, G. M. Di Liberto, M. J. Crosse, and E. C. Lalor, "Electrophysiological Correlates of Semantic Dissimilarity Reflect the Comprehension of Natural, Narrative Speech," *Current Biology*, vol. 28, no. 5, pp. 803–809.e3, 2018.

[16] M. P. Broderick, A. J. Anderson, G. M. Di Liberto, M. J. Crosse, and E. C. Lalor, "Data from: Electrophysiological correlates of semantic dissimilarity reflect the comprehension of natural, narrative speech," Feb. 2019. [Online]. Available: https://doi.org/10.5061/dryad.070jc

[17] G. M. Di Liberto, J. A. O'Sullivan, and E. C. Lalor, "Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing," *Current Biology*, vol. 25, no. 19, pp. 2457–2465, 2015.

[18] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, vol. 88, no. 2, pp. 365–411, 2004.