Multi-subject Resting-State fMRI Data Analysis via Generalized Canonical Correlation Analysis*

Paris A. Karakasis
Dept. of Electrical and Computer Engineering
University of Virginia
Charlottesville, VA
pk4cf@virginia.edu

Athanasios P. Liavas
School of Electrical and Computer Engineering
Technical University of Crete
Chania, Greece
liavas@telecom.tuc.gr

Nicholas D. Sidiropoulos
Dept. of Electrical and Computer Engineering
University of Virginia
Charlottesville, VA
nikos@virginia.edu

Panagiotis G. Simos
School of Medicine
University of Crete
Heraklion, Greece
akis.simos@gmail.com

Efrosyni Papadaki
School of Medicine
University of Crete
Heraklion, Greece
fpapada@otenet.gr

***Abstract*—Functional magnetic resonance imaging (fMRI) is one of the most widespread methods for studying the functionality of the brain. Even at rest, the Blood Oxygen Level Dependent (BOLD) signal reflects systematic fluctuations in the regional brain activity that are attributed to the existence of resting-state brain networks. In many studies, it is assumed that these networks have a common spatially non-overlapping manifestation across subjects, defining a common brain parcellation. In this work, we propose an fMRI data generating model that captures the existence of the common brain parcellation and present a procedure for its estimation. At first, we employ generalized Canonical Correlation Analysis (gCCA) – a well-known statistical method, which can be used for the estimation of a common linear subspace – and recover the subspace that is associated with the common brain parcellation. Then, we obtain an estimate of the common whole-brain parcellation map by solving a semi-orthogonal nonnegative matrix factorization (s-ONMF) problem. We test our theoretical results using both synthetic and real-world fMRI data. Our experimental findings corroborate our theoretical results, rendering our approach a very competitive candidate for multi-subject resting-state whole-brain parcellation.**

***Index Terms*—fMRI, Resting-State, gCCA, MAX-VAR.**

I. Introduction

FUNCTIONAL magnetic resonance imaging (fMRI) is one of the most widespread methods for studying the functionality of the brain. It provides a non-invasive way to measure brain activity, by detecting local changes of the Blood Oxygen Level Dependent (BOLD) signal in the brain, over time. Even at rest, the BOLD signal reflects systematic fluctuations in regional brain activity that are attributed to the existence of resting-state brain networks. A wide range of unsupervised multivariate statistical methods have been applied to fMRI data analysis. Their aim is to provide statistical inference on a whole-brain basis so as to describe brain responses in terms of spatial and temporal patterns. The most commonly used multivariate methods include Principal Component Analysis (PCA) [1], [2], Independent Component Analysis (ICA) [3]–[6], analysis via tensor factorization models [7]–[10], while an extensive overview of fMRI clustering methods and the problem of data driven brain parcellation can be found in [11].

Canonical correlation analysis (CCA) was proposed by Hotelling in [12] and can be considered as a method for the estimation of a linear subspace which is "common" to two sets of random variables [13]. Generalization of CCA to more than two random vectors is a well studied subject [14], [15]. In [16], five formulations of the generalized CCA (gCCA) problem were proposed; all of them boil down to the classical CCA when the number of random vectors is two [17].

A. Problem Definition and Related Work

We focus on multi-subject resting-state fMRI data analysis. Our aim is to extract a data-driven and common, to all subjects, brain parcellation into non-overlapping clusters, without any prior information on the properties of the extracted clusters. In [18], the authors use gCCA to separate common, across subjects, temporal responses to common external stimulation. In contrast, in [19], the assumption of a common spatial response (or spatial map) to the common experimental excitation is adopted while, in [6], under the same assumption, the authors proposed the estimation, via gCCA, of a common spatial subspace that is spanned from the common, across subjects, spatial components, and use it as a preprocessing step before considering the ICA method.

B. Our Contribution

We adopt the assumptions of [19] and [6], regarding the common spatial maps. We compute, via gCCA, an orthonor-

* P. A. Karakasis, A. P. Liavas, P. G. Simos, and E. Papadaki, were partially supported by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship, and Innovation, under the call RESEARCH - CREATE - INNOVATE (project code : T1EΔK-03360)

* N. D. Sidiropoulos was supported in part by NSF ECCS-1807660 and NSF ECCS-1852831.

mal basis for the subspace that is spanned by the common spatial components. We solve a matrix factorization problem with one factor being nonnegative and orthogonal, and compute a common whole-brain parcellation. We compute representative time-series for each cluster and subject, significantly reducing the number of time-series of interest. Finally, we propose a method for determining the number of clusters comprising the brain parcellation. We test our theoretical results using both synthetic and real-world fMRI data.

C. Notation

Scalars are denoted by small letters, vectors by small bold letters, and matrices by capital bold letters, for example, x, $\mathbf{x}$, and $\mathbf{X}$. $\mathbb{R}$ and $\mathbb{R}^{I\times J}$ denote the set of real numbers and the set of $(I \times J)$ real matrices, respectively. Inequality $\mathbf{X} \geq \mathbf{0}$ means that matrix $\mathbf{X}$ has nonnegative elements and $\mathbb{R}_+^{I\times J}$ denotes the set of $(I \times J)$ real matrices with nonnegative elements. $\mathbf{1}_R$ denotes the R-dimensional vector of ones and $\mathbf{I}_R$ denotes the $(R \times R)$ identity matrix. $\|\mathbf{X}\|_2$ and $\|\mathbf{X}\|_F$ denote, respectively, the spectral and the Frobenius norm of matrix $\mathbf{X}$. The transpose and the pseudoinverse of matrix $\mathbf{X}$ are denoted, respectively, by $\mathbf{X}^T$ and $\mathbf{X}^\dagger$. The linear space spanned by the columns of matrix $\mathbf{X}$ is denoted by $\mathrm{col}(\mathbf{X})$. The orthogonal projection onto a linear subspace $\mathcal{S}$ is denoted by $\mathbf{P}_\mathcal{S}$. Finally, we use the Matlab style expressions $\mathbf{X}(:, l)$ and $\mathbf{X}(k, :)$, which denote, respectively, the l-th column and the k-th row of matrix $\mathbf{X}$.

D. Structure

In Section II, we present the data generating model. In Section III, we present a gCCA-based approach for the estimation of the common spatial subspace. We also present a method for the estimation of the dimension of the common spatial subspace. In Section IV, we apply our approach to both synthetic and real-world fMRI data. Finally, in Section V, we conclude the paper.

II. Data Model

Let $\{\mathbf{X}_k\}_{k=1}^K$ be a set of matrices, where $\mathbf{X}_k \in \mathbb{R}^{N\times M}$ denotes the fMRI data of the k-th subject, N denotes the number of voxels, and M denotes the number of time points (in general, $N \gg M$). Let R be a positive integer smaller than M. We adopt the model

$$\mathbf{X}_k = \mathbf{A}\mathbf{S}_k^T + \mathbf{E}_k, \quad k = 1, \ldots, K, \tag{1}$$

where:

1) $\mathbf{A} \in \mathbb{R}_+^{N\times R}$ with $\mathbf{A}^T\mathbf{A} = \mathbf{I}_R$, whose columns are the common, to all subjects, spatial components related with the spontaneous fMRI activity;
2) $\mathbf{S}_k \in \mathbb{R}^{M\times R}$, whose columns are the temporal components, which are associated with the spontaneous fMRI activity and, in general, vary across subjects;
3) $\mathbf{E}_k \in \mathbb{R}^{N\times M}$ denotes the "unmodeled fMRI signal" of the k-th subject and can be considered as (strong) additive noise. We assume that terms $\mathbf{E}_k$ are statistically independent from each other.

Notice that the combination of nonnegativity and orthogonality constraints on spatial factor $\mathbf{A}$ results in the restriction that each row of $\mathbf{A}$ must have at most one nonzero element. Thus, imposing these constraints on $\mathbf{A}$, leads to a data-driven and common, to all subjects, whole-brain parcellation into R non-overlapping clusters. We propose model (1) based on both the work in [3], [5], [6], [19] and the detailed examination of our real-world data.

Our aim is to obtain an accurate estimate of the common spatial term $\mathbf{A}$, leading to a common, to all subjects, precise parcellation of the brain into functionally related clusters.

III. Methods

Our approach for the estimation of the common spatial factor $\mathbf{A}$ and temporal factors $\{\mathbf{S}_k\}_{k=1}^K$ is as follows:

1) we use $\mathbf{X}_k$, for $k = 1, \ldots, K$, and obtain an orthonormal basis for an estimate of the common spatial subspace, $\mathrm{col}(\mathbf{A})$, by solving a gCCA problem;
2) using the solution of the first stage, we obtain an estimate of the common spatial factor $\mathbf{A}$, by solving a matrix factorization problem with one of the factors being nonnegative and orthogonal;
3) using the estimate of factor $\mathbf{A}$, we obtain estimates of the factors $\{\mathbf{S}_k\}_{k=1}^K$.

A. Common Spatial Subspace Estimation via gCCA

We assume that the dimension, R, of the common spatial subspace, $\mathrm{col}(\mathbf{A})$, is known; we shall say more on this important topic later. In order to estimate an orthonormal basis for the common spatial subspace, $\mathrm{col}(\mathbf{A})$, we consider the MAXVAR formulation of the gCCA problem [14]. The associated optimization problem is as follows

$$\min_{\{\mathbf{Q}_k\}_{k=1}^K, \mathbf{G}} \sum_{k=1}^K \|\mathbf{X}_k\mathbf{Q}_k - \mathbf{G}\|_F^2 \ \text{ subject to } \ \mathbf{G}^T\mathbf{G} = \mathbf{I}_R, \tag{2}$$

where $\mathbf{G} \in \mathbb{R}^{N\times R}$ and $\mathbf{Q}_k \in \mathbb{R}^{M\times R}$, for $k = 1, \ldots, K$.

The solution of problem (2) can be computed as follows. Given a matrix $\mathbf{G}$, the optimal $\mathbf{Q}_k$ is given by $\mathbf{Q}_k(\mathbf{G}) = \mathbf{X}_k^\dagger\mathbf{G}$, for $k = 1, \ldots, K$. Substituting this expression into the formulation of problem (2), we obtain the problem

$$\max_{\mathbf{G}} \ \mathrm{Tr}\left(\mathbf{G}^T\mathbf{M}\mathbf{G}\right) \ \text{ subject to } \ \mathbf{G}^T\mathbf{G} = \mathbf{I}_R, \tag{3}$$

where $\mathbf{M} := \sum_{k=1}^K \mathbf{X}_k\mathbf{X}_k^\dagger$. The optimal solution of (3), $\mathbf{G}^o$, is given by [20]

$$\mathbf{G}^o = \mathbf{U}_M\left(:, 1 : R\right), \tag{4}$$

where matrix $\mathbf{U}_M$ emerges from the eigenvalue decomposition of $\mathbf{M}$, i.e. $\mathbf{M} = \mathbf{U}_M\mathbf{\Lambda}_M\mathbf{U}_M^T$. In the noiseless case, where matrices $\mathbf{E}_k = \mathbf{0}$, for $k = 1, \ldots, K$, the solution of problem (2) results into $\mathbf{G}^o$ such that

$$\mathrm{col}(\mathbf{G}^o) = \mathrm{col}(\mathbf{A}). \tag{5}$$

In the presence of noise, terms $\{\mathbf{E}_k\}_{k=1}^K$ are nonzero and, thus, equality (5) is approximate.

B. Estimation of the common spatial factor $\mathbf{A}$

After obtaining the estimate $\mathbf{G}^o$ of an orthonormal basis of the common spatial subspace, we can estimate the common spatial factor $\mathbf{A}$ and temporal factors $\{\mathbf{S}_k\}_{k=1}^K$ by using various approaches. First, notice that, in the absence of noise, relation (5) is equivalent to $\mathbf{G}^o = \mathbf{A}\mathbf{B}^T$, where $\mathbf{B} \in \mathbb{R}^{R\times R}$. In the presence of noise, we may estimate $\mathbf{A}$ by solving the semi-Orthogonal Nonnegative Matrix Factorization (s-ONMF) problem [21]

$$\min_{\mathbf{A},\,\mathbf{B}} \left\|\mathbf{G}^o - \mathbf{A}\mathbf{B}^T\right\|_F^2, \text{ subject to } \mathbf{A}^T\mathbf{A} = \mathbf{I}_R, \ \mathbf{A} \geq \mathbf{0}. \quad (6)$$

This is an extensively-studied problem which can be attacked in many ways. In our experiments, we use the Penalty Method. That is, for an increasing sequence of positive penalty coefficients $\{c_k\}$, we solve subproblems of the form [22]

$$\min_{\mathbf{A},\,\mathbf{B}} \left\|\mathbf{G}^o - \mathbf{A}\mathbf{B}^T\right\|_F^2 + c_k \text{ trace}\left(\mathbf{Q}\mathbf{A}^T\mathbf{A}\right), \text{ subject to } \mathbf{A} \geq \mathbf{0},$$

where $\mathbf{Q} = \mathbf{1}_R\mathbf{1}_R^T - \mathbf{I}_R$. Given the optimal matrix, $\mathbf{A}^o$, we obtain estimates of factors $\{\mathbf{S}_k\}_{k=1}^K$ by solving least-squares problems, i.e. $\mathbf{S}_k^o = \mathbf{X}_k^T\mathbf{A}^o$.

C. On the Dimension of the Common Spatial Subspace

In Subsection III-A, we assumed that the true dimension, R, of the common spatial subspace, $\text{col}(\mathbf{A})$, is known. Of course, in general, the value of R is unknown, thus, we must estimate it from the data. In the sequel, we provide a procedure which gives us very useful information about the value of R.

Let the assumed dimension of the common spatial subspace be $\hat{R} = R$ and $\{\mathcal{K}_1, \mathcal{K}_2\}$ be a random partition of the set of the subjects $\{1, \ldots, K\}$. In the noiseless case, solving the problem (2) twice, for $k \in \mathcal{K}_1$ and $k \in \mathcal{K}_2$, results into orthonormal bases $\mathbf{G}_1^o$ and $\mathbf{G}_2^o$, with $\text{col}(\mathbf{G}_1^o) = \text{col}(\mathbf{G}_2^o)$. If we start adding noise and repeat the procedure, then we will obtain $\text{col}(\mathbf{G}_1^o)$ and $\text{col}(\mathbf{G}_2^o)$ which will be "close" to each other. One way to measure the distance between linear subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$ is to compute their gap, defined as [23, p. 93]

$$\rho_{g,2}\left(\mathcal{S}_1, \mathcal{S}_2\right) := \left\|\mathbf{P}_{\mathcal{S}_1} - \mathbf{P}_{\mathcal{S}_2}\right\|_2. \quad (7)$$

If $\hat{R} = R$ and $\|\mathbf{E}_k\|_2 = O(\epsilon)$, for $k = 1, \ldots, K$, where ϵ is a small positive number, then we expect that

$$\left\|\mathbf{P}_{\text{col}(\mathbf{G}_1^o)} - \mathbf{P}_{\text{col}(\mathbf{G}_2^o)}\right\|_2 = O(\epsilon). \quad (8)$$

If $\hat{R} > R$, then, by solving (2), besides the R-dimensional common subspace, $\text{col}(\mathbf{A})$, we shall try to model "common" noise subspace. Since the noise terms $\mathbf{E}_k$ are independent across subjects and $N \gg M$, we do *not* expect to find any common noise subspace in the two data sets associated with $\mathcal{K}_1$ and $\mathcal{K}_2$. Thus, in this case, we expect that

$$\left\|\mathbf{P}_{\text{col}(\mathbf{G}_1^o)} - \mathbf{P}_{\text{col}(\mathbf{G}_2^o)}\right\|_2 \approx 1. \quad (9)$$

Finally, if $\hat{R} < R$ and the rank-one terms that constitute the products $\mathbf{A}\mathbf{S}_k^T$ are of almost "equal" strength, then we expect that

$$O(\epsilon) \leq \left\|\mathbf{P}_{\text{col}(\mathbf{G}_1^o)} - \mathbf{P}_{\text{col}(\mathbf{G}_2^o)}\right\|_2 \lessapprox 1, \quad (10)$$

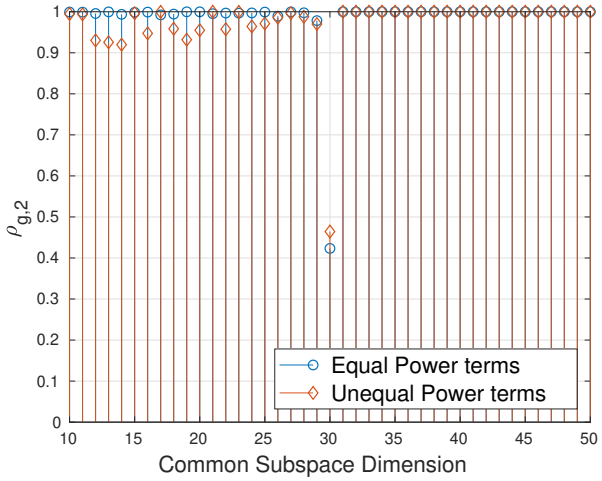

Fig. 1: Gap of $\text{col}(\mathbf{G}_1^o)$ and $\text{col}(\mathbf{G}_2^o)$ as a function of the subspace dimension, for SNR $= -5$dB. Blue color: rank-one terms of the products $\mathbf{A}\mathbf{S}_k^T$ have (on average) the same Frobenius norm (on average). Red color: rank-one terms have (on average) different Frobenius norms. The true common subspace dimension is $R = 30$.

because $\text{col}(\mathbf{G}_1^o)$ and $\text{col}(\mathbf{G}_2^o)$ will "randomly" capture $\hat{R}$ out of R dimensions of the common spatial subspace.

Thus, the gap between $\text{col}(\mathbf{G}_1^o)$ and $\text{col}(\mathbf{G}_2^o)$ provides valuable information about the dimension of the common subspace, $\text{col}(\mathbf{A})$. Accurate expressions for the gap lie beyond the scope of this manuscript, require tools from matrix perturbation theory, and pose stringent assumptions on the size of the noise, which may not be fulfilled in our case. We shall check the usefulness of our claims in Subsection IV.

IV. Experiments

A. Synthetic Data

In this subsection, we test the effectiveness of our approach using synthetic data. More specifically, we generate random data according to the model

$$\mathbf{X}_k = \mathbf{A}\mathbf{S}_k^T + \beta\mathbf{E}_k, \quad (11)$$

where $\mathbf{A}$ is generated as follows: for each row i, we select uniformly at random a column j, and we draw element $\mathbf{A}_{i,j}$ from $\mathcal{U}[0,1]$. Then, we normalize the columns of $\mathbf{A}$ to meet the orthonormality constraints $\mathbf{A}^T\mathbf{A} = \mathbf{I}_R$. Matrices $\mathbf{E}_k$ and columns of matrices $\mathbf{S}_k$, $\mathbf{S}_k(:, r)$, for $k = 1, \ldots, K$ and $r = 1, \ldots, R$, have i.i.d $\mathcal{N}(0,1)$ and $\mathcal{N}(0, \sigma_r^2)$ elements, respectively, where σ_r^2 may vary across r, depending on the scenario we consider. The Signal-to-Noise Ratio (SNR), defined as

$$\text{SNR} := \frac{\sum_{k=1}^K \left\|\mathbf{A}\mathbf{S}_k^T\right\|_F^2}{\beta^2 \sum_{k=1}^K \left\|\mathbf{E}_k\right\|_F^2}, \quad (12)$$

is determined by the scalar factor β. In our experiments, we set $N = 10^5$, $M = 100$, $K = 25$, and $R = 30$.

In order to illustrate the usefulness of our dimension determination criterion, in Fig.1, we plot the gap between $\text{col}(\mathbf{G}_1^o)$ and $\text{col}(\mathbf{G}_2^o)$ as a function of the common spatial subspace

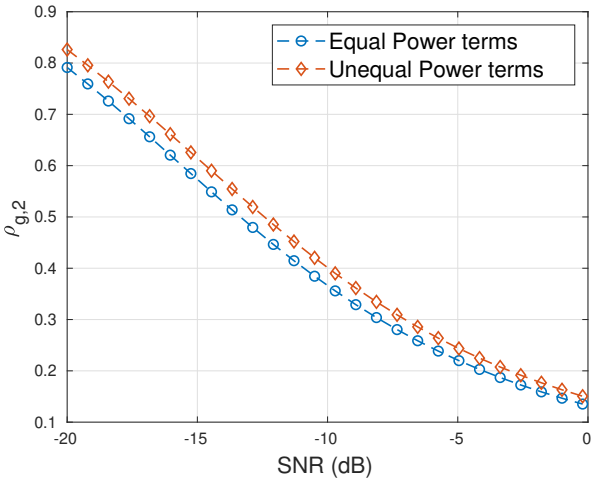

Fig. 2: Gap between $\mathrm{col}(\mathbf{A})$ and $\mathrm{col}(\mathbf{G}^o)$ versus SNR.

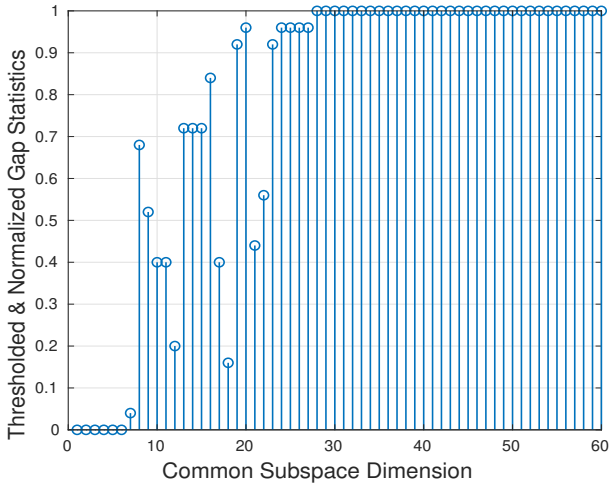

Fig. 3: Common subspace dimension estimation using real-world fMRI data.

dimension, for $\mathtt{SNR} = -5$dB. We consider two scenarios. In the first, the rank-one terms that constitute the products $\mathbf{A}\mathbf{S}_k^T$ are of equal power (we set $\sigma_r^2 = 1$ for $r = 1, \ldots, R$), while in the second, they are of unequal power, with $0.5 \leq \sigma_r^2 \leq 1$, for $r = 1, \ldots, R$. We observe that, in both cases, *all* values of the common spatial subspace dimension that are larger than the true value $R = 30$ lead to gaps that are practically equal to 1. Thus, we are able to determine a critical value for the common subspace dimension.

In Fig. 2, we plot the gap between $\mathrm{col}(\mathbf{A})$ and $\mathrm{col}(\mathbf{G}^o)$, as a function of the SNR, for the true common spatial subspace dimension, $R = 30$. We observe that, in both scenarios, we attain accurate estimates for SNR higher than -5 dB.

B. Real World Data

In the sequel, we test our approach using real-world resting-state fMRI data. The data was acquired during the period 2015-2016 from 31 healthy subjects and was recorded in the MRI Unit, University Hospital of Heraklion. The hospital review board approved the study and the procedure was thoroughly explained to all patients and volunteers, who signed informed consent before undergoing MRI.

Brain MRI examinations were performed on a clinical, upgraded $1.5T$ Siemens Vision/Sonata scanner (Erlangen, Germany), Gradient strength: $45\,mT/m$, Gradient slew rate: $200\,mT/m/ms$ and a standard four-channel head array coil (minimum voxel dimensions: $70\,\mu m \times 70\,\mu m \times 300\,\mu m$). Resting-state functional MRI (rs-fMRI) was derived from a $T2*$-weighted, fat-saturated $2D$-FID-EPI sequence with repetition time (TR) $3500\,ms$, echo time (TE) $50\,ms$, field of view (FOV) $192 \times 192 \times 108\,(x,\, y,\, z)$, and acquisition voxel size $3 \times 3 \times 3\,mm$. Whole brain scans consist of 36 transverse slices with 3.0-mm slice thickness and no interslice gap. Each BOLD time–series consists of 150 dynamic volumes, however, the first 5 volumes were discarded as is usual in fMRI studies. The fMRI images were smoothed, normalized, and co-registered to the MNI space. The time-series were detrended (subtraction of the mean and the linear trend) prior to the application of the gCCA method.

In Fig. 3, we depict the results that emerged by the following procedure. We consider 25 random partitions $\mathcal{K}_1$ and $\mathcal{K}_2$ of the set of the subjects $\{1, \ldots, 31\}$. For each partition and each common subspace dimension, we estimate the orthonormal bases $\mathbf{G}_1^o$ and $\mathbf{G}_2^o$ and compute their gap. We use the threshold value 0.95, and set the gap values that are larger than the threshold equal to 1, and those that are smaller than the threshold equal to 0. For each dimension, we sum the thresholded gap values we obtained from the 25 partitions and divide the sum by 25. We call the resulting values as "Thresholded and Normalized Gap Statistics." The resulting plot indicates that a good estimate is $R = 27$.

In Fig. 4, we depict the clusters computed by our approach, for $R = 27$. In Fig. 5, we illustrate how the clusters change when we increase the value of R from 26 to 27. At the top, we plot the maximum correlation coefficient between each column of $\mathbf{A}_{27}$ with the columns of $\mathbf{A}_{26}$; at the bottom, we depict the corresponding pairing. We observe that most of the columns of $\mathbf{A}_{27}$ are highly correlated with columns of $\mathbf{A}_{26}$. The smallest correlation values at the top appear at positions 14 and 16; looking at the bottom, we see that both these columns of $\mathbf{A}_{27}$ are associated with column 18 of $\mathbf{A}_{26}$, meaning that the respective cluster for $R = 26$ has been partitioned into two clusters for $R = 27$.

An interesting future topic is the comparison of our method with other methods, for example, [24].

V. Conclusion

We considered the problem of multi-subject resting-state fMRI analysis. We derived a whole-brain parcellation map, by applying gCCA and solving a semi-nonnegative orthogonal matrix factorization problem. Our approach has been proven effective with both synthetic and real-world data.

Acknowledgment

The authors would like to thank Margarita-Antonia Psychountaki for help with some of the experiments.

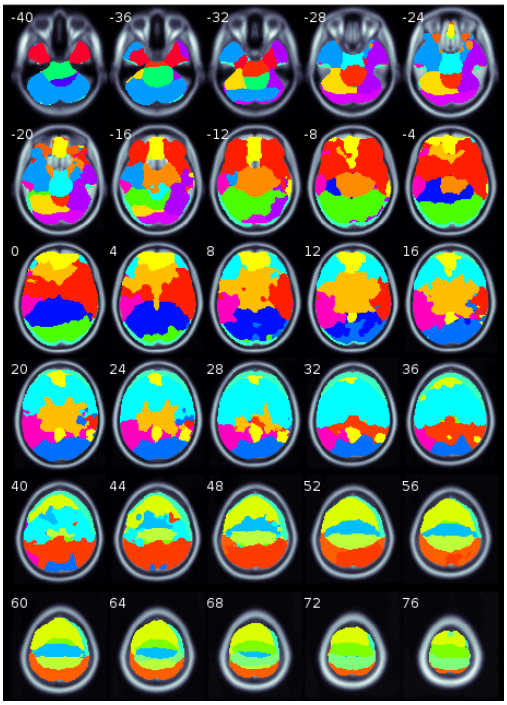

Fig. 4: The gCCA-based whole-brain parcellation for $R = 27$.

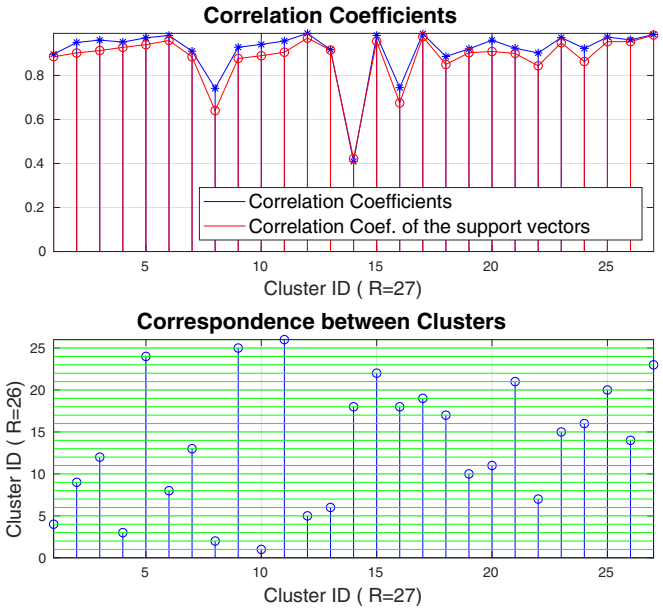

Fig. 5: Top: correlation coefficients between pairs of clusters for $R = 26$ and $R = 27$. Bottom: the cluster correspondence.

REFERENCES

[1] A. H. Andersen, D. M. Gash, and M. J. Avison, "Principal component analysis of the dynamic response measured by fMRI: a generalized linear systems framework," *Magnetic Resonance Imaging*, vol. 17, no. 6, pp. 795–815, 1999.

[2] L. K. Hansen, J. Larsen, F. Å. Nielsen, S. C. Strother, E. Rostrup, R. Savoy, N. Lange, J. Sidtis, C. Svarer, and O. B. Paulson, "Generalizable patterns in neuroimaging: How many principal components?," *NeuroImage*, vol. 9, no. 5, pp. 534–544, 1999.

[3] V. Calhoun, T. Adali, G. Pearlson, and J. Pekar, "Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms," *Human brain mapping*, vol. 13, no. 1, pp. 43–53, 2001.

[4] V. D. Calhoun, T. Adali, L. K. Hansen, J. Larsen, and J. J. Pekar, "ICA of functional MRI data: an overview," in *in Proceedings of the International Workshop on Independent Component Analysis and Blind Signal Separation*, Citeseer, 2003.

[5] E. B. Erhardt, S. Rachakonda, E. J. Bedrick, E. A. Allen, T. Adali, and V. D. Calhoun, "Comparison of multi-subject ICA methods for analysis of fMRI data," *Human brain mapping*, vol. 32, no. 12, pp. 2075–2095, 2011.

[6] G. Varoquaux, S. Sadaghiani, P. Pinel, A. Kleinschmidt, J.-B. Poline, and B. Thirion, "A group model for stable multi-subject ica on fmri datasets," *Neuroimage*, vol. 51, no. 1, pp. 288–299, 2010.

[7] A. H. Andersen and W. S. Rayens, "Structure-seeking multilinear methods for the analysis of fMRI data," *NeuroImage*, vol. 22, no. 2, pp. 728–739, 2004.

[8] C. F. Beckmann and S. M. Smith, "Tensorial extensions of independent component analysis for multisubject fMRI analysis," *Neuroimage*, vol. 25, no. 1, pp. 294–311, 2005.

[9] C. Chatzichristos, E. Kofidis, Y. Kopsinis, M. M. Moreno, and S. Theodoridis, "Higher-order block term decomposition for spatially folded fMRI data," in *International Conference on Latent Variable Analysis and Signal Separation*, pp. 3–15, Springer, 2017.

[10] K. H. Madsen, N. W. Churchill, and M. Mørup, "Quantifying functional connectivity in multi-subject fMRI data using component models," *Human brain mapping*, vol. 38, no. 2, pp. 882–899, 2017.

[11] B. Thirion, G. Varoquaux, E. Dohmatob, and J.-B. Poline, "Which fmri clustering gives good brain parcellations?," *Frontiers in neuroscience*, vol. 8, p. 167, 2014.

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

[13] M. S. Ibrahim and N. D. Sidiropoulos, "Cell-edge interferometry: Reliable detection of unknown cell-edge users via canonical correlation analysis," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, IEEE, 2019.

[14] P. Horst, "Generalized canonical correlations and their applications to experimental data," *Journal of Clinical Psychology*, vol. 17, no. 4, pp. 331–347, 1961.

[15] J. D. Carroll, "Generalization of canonical correlation analysis to three or more sets of variables," in *Proceedings of the 76th annual convention of the American Psychological Association*, vol. 3, pp. 227–228, 1968.

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

[17] N. A. Asendorf, "Informative data fusion: Beyond canonical correlation analysis," 2015.

[18] Y.-O. Li, W. Wang, T. Adali, and V. D. Calhoun, "CCA for joint blind source separation of multiple datasets with application to group fMRI analysis," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1837–1840, IEEE, 2008.

[19] B. Afshin-Pour, G.-A. Hossein-Zadeh, S. C. Strother, and H. Soltanian-Zadeh, "Enhancing reproducibility of fMRI statistical maps using generalized canonical correlation analysis in NPAIRS framework," *NeuroImage*, vol. 60, no. 4, pp. 1970–1981, 2012.

[20] X. Fu, K. Huang, M. Hong, N. D. Sidiropoulos, and A. M.-C. So, "Scalable and flexible multiview MAX-VAR canonical correlation analysis," *IEEE Transactions on Signal Processing*, vol. 65, no. 16, pp. 4150–4165, 2017.

[21] C. H. Ding, T. Li, and M. I. Jordan, "Convex and semi-nonnegative matrix factorizations," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 1, pp. 45–55, 2008.

[22] B. Li, G. Zhou, and A. Cichocki, "Two efficient algorithms for approximately orthogonal nonnegative matrix factorization," *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 843–846, 2014.

[23] G. Stewart and J. Sun, "Computer science and scientific computing. matrix perturbation theory," 1990.

[24] B. Thomas Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, *et al.*, "The organization of the human cerebral cortex estimated by intrinsic functional connectivity," *Journal of neurophysiology*, vol. 106, no. 3, pp. 1125–1165, 2011.