

Dominant Component Analysis of Electrophysiological Connectivity Networks

Yasser Ghanbari¹, Luke Bloy¹, Kayhan Batmanghelich¹,
Timothy P.L. Roberts^{2,*}, and Ragini Verma¹

¹ Section of Biomedical Image Analysis, University of Pennsylvania, Philadelphia, PA
{Yasser.Ghanbari,Luke.Bloy,Nematollah.Batmanghelich,
Ragini.Verma}@uphs.upenn.edu

² Lurie Family Foundation's MEG Imaging Center, Department of Radiology,
Children's Hospital of Philadelphia, Philadelphia, PA, USA
robertstim@email.chop.edu

Abstract. Connectivity matrices obtained from various modalities (DTI, MEG and fMRI) provide a unique insight into brain processes. Their high dimensionality necessitates the development of methods for population-based statistics, in the face of small sample sizes. In this paper, we present such a method applicable to functional connectivity networks, based on identifying the basis of dominant connectivity components that characterize the patterns of brain pathology and population variation. Projection of individual connectivity matrices into this basis allows for dimensionality reduction, facilitating subsequent statistical analysis. We find dominant components for a collection of connectivity matrices by using the projective non-negative component analysis technique which ensures that the components have non-negative elements and are non-negatively combined to obtain individual subject networks, facilitating interpretation. We demonstrate the feasibility of our novel framework by applying it to simulated connectivity matrices as well as to a clinical study using connectivity matrices derived from resting state magnetoencephalography (MEG) data in a population of subjects diagnosed with autism spectrum disorder (ASD).

1 Introduction

Although the neurological origin of many brain disorders is still unknown, computational techniques applied to neuroimaging data help unveil the underlying functional or structural differences between patient and typically developing (TD) populations. Many brain disorders such as ASD and schizophrenia are believed to be due, in part, to altered brain connectivity [1–4]. Hence, connectivity analyses of brain function has received considerable attention as an indicator of or biomarker for brain disorders. Such studies have investigated the functional or structural connectivity networks using fMRI, DTI, EEG, or MEG recordings

* The authors would like to acknowledge support from the NIH grants: MH092862 and DC008871.

by calculating correlation, synchronization, or mutual information measures and interpreting abnormalities such as deficient long-range connections [1, 3].

A number of analysis approaches exists to utilize connectivity methods to investigate the underlying brain processes occurring during rest as well as under specific task conditions. A successful analysis methodology must possess a means of identifying relevant sub-networks providing an interpretable representation of the brain activity, while also facilitating the statistical analysis able to describe how this representation is affected by disease. The approach taken here is the decomposition of connectivity matrices into dominant components while enforcing positivity of both the components and coefficients. Such a decomposition maintains the interpretation of each component as a connectivity matrix and the coefficients as activations of those networks while providing a succinct low dimensional representation of the population amenable to statistical analysis. The traditional approaches, principal and independent components analyses (PCA and ICA), used for investigating brain networks [5, 6] provide dimensionality reduction but lack the necessary constraints that provide a physiologic interpretability of the decomposition.

In this paper, we present the projective non-negative component analysis to identify the dominant functional connectivity networks in a population. These networks form a basis of variation in the population that could be due to differences in age, pathology, gender, etc. The non-negativity constraint on both components and the coefficients allows the interpretation of a difference in the coefficient of a particular component as a change in the degree to which it is activated. Additionally, the connectivity components are nearly orthogonal. The orthogonality of components with non-negative elements happens only when they do not share non-zero dimensions, leading to sparsity [7, 8]. Such unique analysis enables us to obtain the dominant networks, thereby providing a global view of the brain processes and how they are affected by disease.

The applicability of the proposed method is demonstrated in simulated connectivity matrices while comparing it with an alternate form of basis decomposition. While the method is generalizable to any type of connectivity matrix from alternate modalities, in this work we apply it to electrophysiological functional connectivity networks computed from resting-state MEG data by using the synchronization likelihood (SL) analysis [9]. This novel approach utilizes the high temporal resolution of MEG data, allowing an accurate characterization of the functional coupling between different areas of the brain. When coupled with a suitable analysis framework, such as the proposed, MEG based connectivity promises valuable insights into the temporal dynamics of brain states that is unavailable from other modalities as well as how they are affected by pathology.

2 Methods

2.1 Projective Non-Negative Component Analysis (PNCA)

We hypothesize that each connectivity matrix obtained from recording the brain activity of a subject is a linear combination of several fundamental connectivity

matrices called connectivity components. Due to the symmetry of connectivity matrices, a vector of all elements of the upper triangular part of any connectivity matrix is considered as a representative of that matrix, and is used as an observation vector \mathbf{y}_i for the corresponding subject i . To compute the connectivity components whose mixture approximately constructs the observed connectivity matrices, a matrix factorization model is used as follows.

$$\mathbf{Y} \approx \mathbf{W}\Phi, \tag{1}$$

where columns of $\mathbf{Y}_{p \times q}$, i.e. \mathbf{y}_i ($1 \leq i \leq q$), are the SL connectivity matrix representatives, and columns of $\mathbf{W}_{p \times r}$, i.e. \mathbf{w}_j ($1 \leq j \leq r$), are representative of the basis connectivity components, i.e. the upper triangular elements of the matrix of the corresponding connectivity component. These components (\mathbf{w}_j) are then mixed by the elements of each column of the loading matrix $\Phi_{r \times q}$ to approximate the corresponding column of \mathbf{Y} [10, 8].

Assuming that Φ is the projection of \mathbf{Y} onto \mathbf{W} , i.e. $\Phi = \mathbf{W}^T \mathbf{Y}$, the non-negativity constraint on the elements of \mathbf{W} and Φ makes our non-negative component analysis an optimization problem of minimizing the cost function $F(\mathbf{W}) = \|\mathbf{Y} - \mathbf{W}\mathbf{W}^T \mathbf{Y}\|^2$ with respect to \mathbf{W} , where $\|\cdot\|$ represents the matrix norm. Considering the Frobenius norm, the optimization problem can be denoted by [10]

$$\min_{\mathbf{W} \geq 0} F_{fro}(\mathbf{W}) = \min_{\mathbf{W} \geq 0} trace \left\{ (\mathbf{Y} - \mathbf{W}\mathbf{W}^T \mathbf{Y}) (\mathbf{Y} - \mathbf{W}\mathbf{W}^T \mathbf{Y})^T \right\}. \tag{2}$$

The above cost function can be minimized by a gradient descent approach, i.e. updating $W_{i,j} = W_{i,j} - \eta_{i,j} \frac{\partial F}{\partial W_{i,j}}$ with a positive step-size $\eta_{i,j}$ where

$$\frac{\partial F}{\partial W_{i,j}} = -4 (\mathbf{Y}\mathbf{Y}^T \mathbf{W})_{i,j} + 2 (\mathbf{W}\mathbf{W}^T \mathbf{Y}\mathbf{Y}^T \mathbf{W})_{i,j} + 2 (\mathbf{Y}\mathbf{Y}^T \mathbf{W}\mathbf{W}^T \mathbf{W})_{i,j}. \tag{3}$$

The gradient descent updating, as stated above, does not guarantee to keep the elements of \mathbf{W} non-negative. However, due to the positivity of the elements of \mathbf{Y} and by applying positive initialization to \mathbf{W} , our non-negativity constraint is guaranteed by having the step-size as follows

$$\eta_{i,j} = \frac{\frac{1}{2} W_{i,j}}{(\mathbf{W}\mathbf{W}^T \mathbf{Y}\mathbf{Y}^T \mathbf{W})_{i,j} + (\mathbf{Y}\mathbf{Y}^T \mathbf{W}\mathbf{W}^T \mathbf{W})_{i,j}}, \tag{4}$$

which results in the following multiplicative updating procedure [10]

$$W_{i,j} = W_{i,j} \frac{2 (\mathbf{Y}\mathbf{Y}^T \mathbf{W})_{i,j}}{(\mathbf{W}\mathbf{W}^T \mathbf{Y}\mathbf{Y}^T \mathbf{W})_{i,j} + (\mathbf{Y}\mathbf{Y}^T \mathbf{W}\mathbf{W}^T \mathbf{W})_{i,j}}. \tag{5}$$

For stability of the convergence, at each iteration, \mathbf{W} is normalized by $\mathbf{W} = \frac{\mathbf{W}}{\|\mathbf{W}\|_2}$. Starting by initial random positive elements on \mathbf{W} , the iterative procedure will converge to the desired $\mathbf{W} \geq 0$ whose columns are PNCA. Each obtained component \mathbf{w}_j (the j^{th} column of \mathbf{W}) is then normalized by its norm $\|\mathbf{w}_j\|$,

and correspondingly the coefficients, i.e. elements of the matrix $\Phi = [\varphi_{ji}]$, are adjusted as $\varphi_{ji} = \frac{\varphi_{ji}}{\|\mathbf{w}_j\|} : 1 < i < q$.

The resulting non-negative coefficients are an indicator of the weight of the corresponding component in reconstructing the original connectivity matrices. Therefore, we rank each component \mathbf{w}_j based on the average of corresponding coefficients, i.e. $\frac{1}{q} \sum_{i=1}^q \varphi_{ji}$.

2.2 Group PNCA Model for SL Networks

As stated by (1), the q connectivity observations, i.e. $\mathbf{y}_i : 1 \leq i \leq q$, in the matrix \mathbf{Y} are approximated by

$$[\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_q] \approx [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r] \begin{bmatrix} \varphi_{11} & \varphi_{12} & \dots & \varphi_{1q} \\ \vdots & \vdots & & \vdots \\ \varphi_{r1} & \varphi_{r2} & \dots & \varphi_{rq} \end{bmatrix}. \quad (6)$$

Each observation vector per subject i is thus, approximately reconstructed by

$$\mathbf{y}_i \approx \sum_{j=1}^r \varphi_{ji} \mathbf{w}_j = \sum_{j=1}^r (\mathbf{w}_j^T \mathbf{y}_i) \mathbf{w}_j ; 1 \leq i \leq q. \quad (7)$$

Let us suppose, with no loss of generality, that the first q_1 elements are from the first group (e.g. population of ASD) and the remaining from the second group (e.g. TD group). Thereby, the role of each component \mathbf{w}_j in reconstructing the corresponding connectivity vector in the first group, i.e. $\mathbf{y}_i : 1 \leq i \leq q_1$, is characterized by the corresponding coefficients φ_{ji} ; and so forth for the second group. Therefore, the statistical significance between the set of $\{\varphi_{ji} : 1 \leq i \leq q_1\}$ and $\{\varphi_{ji} : q_1 + 1 \leq i \leq q\}$ describes the importance of the corresponding basis connectivity component \mathbf{w}_j in differentiating the two groups.

2.3 Synchronization Likelihood (SL) Connectivity Networks

Time-frequency synchronization likelihood assumes that two signals are highly synchronized if a pattern of one signal repeats itself at certain time instants within a time period while the other signal repeats itself at those same instants [9]. Such signal patterns at a time instant t_i of channel k can be represented by an embedding vector $\mathbf{x}_{k,t_i} = [x_k(t_i), x_k(t_{i+l}), \dots, x_k(t_{i+(m-1)l})]$ where l is the lag and m is the length of the embedding vector. l and m are typically set to $l = \frac{f_s}{3h_f}$ and $m = \frac{3h_f}{l_f} + 1$ where f_s is the sampling frequency, and h_f and l_f are the high and low frequency contents of the signal, respectively [9]. At each time instant t_i , the Euclidean distance is measured between the reference embedding vector \mathbf{x}_{k,t_i} and the set of all other embedding vectors at times t_j , i.e. \mathbf{x}_{k,t_j} , where t_j lies in the range $t_i - \frac{t_{w_2}}{2} < t_j < t_i - \frac{t_{w_1}}{2}$ or $t_i + \frac{t_{w_1}}{2} < t_j < t_i + \frac{t_{w_2}}{2}$ (in this work t_{w_2} was set to 10 sec and $t_{w_1} = \frac{2l(m-1)}{f_s}$). Then, n_{ref} nearest embedding vectors \mathbf{x}_{k,t_j} are retained [9].

This procedure is conducted for each channel k and each time instants t_i . Then, the SL between channel k_1 and channel k_2 at time instant t_i is the number of simultaneous recurrences in the two channels divided by the total number of recurrences, i.e. $SL_{t_i} = n_{k_1 k_2} / n_{ref}$. The synchronization likelihood at all the time instants t_i are then averaged yielding the final SL between the two channels.

3 Results

3.1 Simulation Experiments

In order to demonstrate the effectiveness of the proposed method of PNCA in identifying the true underlying connectivity components, we apply our method to simulated connectivity matrices. We compare this with ICA which is the most widely-adopted method for similar purposes. The simulation is performed by using random non-negative numbers as the elements of three 10×10 simulated symmetric SL matrices plus a small background random (non-negative) noise. Ten linear mixtures of the simulated connectivity matrices are composed by a random mixing matrix with non-negative elements. The upper triangular part (excluding the diagonal) of the 10 connectivity matrices are vectorized to form the 10 columns of $\mathbf{Y}_{45 \times 10}$. We apply the fast ICA algorithm [11] as well as the proposed technique in Sect. 2.1 to solve for $r = 3$ normalized components as columns of $\mathbf{W}_{45 \times 3}$. For visualization, the SL matrices and the elements of the resulting components from fast ICA and PNCA algorithms are displayed by grayscale images. The results are shown in Fig. 1. It can be seen that while our method produces the right components (as compared to the original components in Fig. 1(a)), ICA is unable to resolve the components correctly. This is due to the fact that ICA forces the components to be statistically independent (i.e. with covariance of identity) whereas PNCA forces the components to be non-negative and orthogonal which leads to localized components due to sparsity. Please note that the resulting components are not ranked and ordered in Fig. 1.

3.2 PNCA on MEG SL Connectivity Networks

Dataset and preprocessing. Our dataset consisted of 48 children subjects (26 ASD's and 22 TD's) aged 6-15 years (mean=10.1, SD=2.3 in ASD, and mean=11.0, SD=2.5 in TD). A standard t-test showed that the populations' age difference was not statistically significant ($p - value > 0.19$). All subjects reported are free from medication use. Resting-state data were collected by a 275-channel MEG system, 274 channels being effective at the time of recording, with three head-position indicator coils to monitor head motion and ensure immobility. After a band-pass filter (0.03-150 Hz), MEG signals were digitized at 1200 Hz and downsampled offline to 500 Hz. Eye-blink artifacts were removed by learning the pattern of eye blinks and deleting that component from the data.

The recordings are then used to calculate the 274×274 SL matrices for delta (0.5–4 Hz) frequency band using a fourth order Butterworth bandpass filter. This

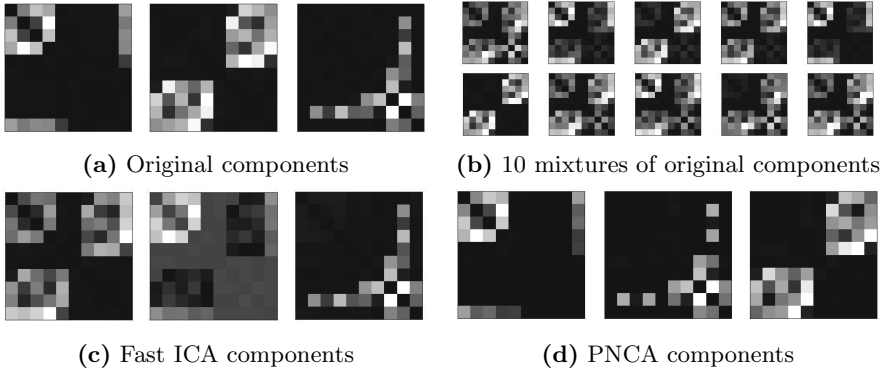


Fig. 1. The results of fast ICA and PNCA on the simulated SL matrices. (a) The simulated SL matrices as the original components. (b) 10 linear mixtures of the components by random weightings. (c) The solution of fast ICA. (d) The solution of PNCA.

could be computed in any frequency band, but ASD studies, e.g. [1], have shown delta band anomalies and hence we concentrate on the delta band. Therefore, a total of 26 and 22 SL matrices are obtained for 26 ASD and 22 TD subjects, respectively. The upper triangular (excluding the diagonal) elements of each SL matrix are concatenated to make a vector of $\frac{274 \times 273}{2} = 37401$ elements which form one column of $\mathbf{Y}_{37401 \times 48}$.

Component Analysis. In order to determine the number of components used in the PNCA decomposition, we separately apply PCA to the aggregated connectivity vectors of populations of ASD, TD, and pooled ASD and TD (needed for a joint statistical analysis). The PCA findings indicated that a decomposition consisting of five ($r = 5$) components would account for 93% to 95% of the total variance. The PNCA is applied to the three cases of ASD ($\mathbf{Y}_{37401 \times 26}$), TD ($\mathbf{Y}_{37401 \times 22}$), and pooled ASD and TD ($\mathbf{Y}_{37401 \times 48}$), and the model of (1) is solved for five components ($\mathbf{W}_{37401 \times 5}$) for each of the three population cases. The resulting 37401-length connectivity components at each column of \mathbf{W} are then used to form the corresponding 274×274 symmetric connectivity matrices. Figure 2 shows the resulting five connectivity components on the MEG sensor map. These components are ranked here (left to right) based on the descending average of their corresponding coefficients in each of the three population cases. The averages of the coefficients are given in Table 1.

A statistical group analysis, as described in Sect. 2.2, was performed over the resulting PNCA coefficients, i.e. φ_{ji} . The two-sample t-test is applied to the coefficients of each normalized connectivity component from the pooled ASD-TD and the p -values and t -values are given in Table 1.

We see that the average of the first component weights (φ_{1i}) are statistically smaller in constructing the SL connectivity matrices of the ASD group compared to the TD group. Interestingly, this is the most dominate component, ranked by the average of the corresponding coefficients, in the pooled

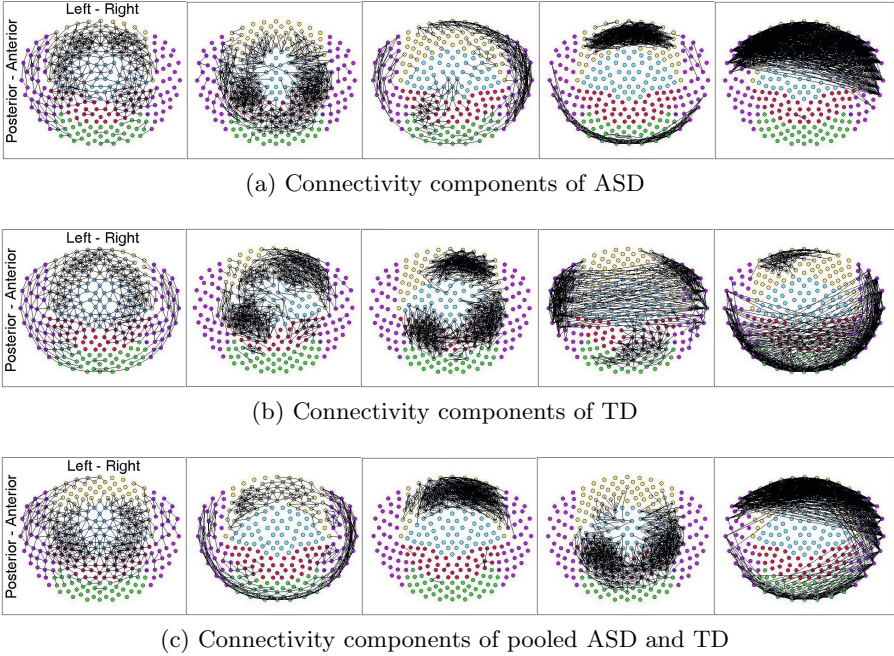


Fig. 2. PNCA connectivity components plotted on the MEG sensor map and sorted based on their average coefficients

Table 1. PNCA component ranking and statistical group analysis

Component No.	ASD Average Coeff's	TD Average Coeff's	Pooled ASD-TD Average Coeff's	ASD-TD group p -value	ASD-TD group t -value
1	10.2	11.4	9.7	0.02	-2.3
2	6.2	6.4	7.0	0.32	-1.0
3	6.0	4.8	5.7	0.75	+0.3
4	4.6	4.6	5.0	0.28	-1.1
5	2.0	2.8	2.8	0.23	+1.2

population, and is quite similar to the most dominant components determined from the separate ASD and TD populations. This component can be interpreted as the default connectivity network observed in the resting brain. It is also notable that the presence of the fourth and fifth components of ASD indicates strong short range frontal connectivity diminished in TD, while the fourth component in TD population, with clear long range connectivity, is not found in the ASD set of components. Together these support intact default connectivity in ASD with evidence for diminished long range and enhanced short-range connectivity in ASD; a finding consistent with other connectivity analysis investigations in autism [1].

4 Conclusion

We have presented a non-negative component analysis technique for learning localized and part-based sparse components of non-negative connectivity matrices. The algorithm is based on non-negative projections which produces non-negative bases and coefficients by a gradient descent approach minimizing a Frobenius norm of the reconstruction matrix error. We applied it to the simulated connectivity matrices which showed more accurate component findings compared to the well-known ICA technique. The proposed method is then applied to the novel problem of investigating MEG derived SL connectivity matrices, within a group study of ASD. The projections of the connectivity elements onto the components revealed statistically significant differences between how the ASD and TD functional connectivity matrices are composed of their fundamental connectivity components. The presented technique represents a framework, in principle capable of handling other types of functional or structural connectivity networks from any modality for statistical group analysis or feature selection.

References

1. Barttfeld, P., et al.: A big-world network in asd: Dynamical connectivity analysis reflects a deficit in long-range connections and an excess of short-range connections. *Neuropsychologia* 49, 254–263 (2011)
2. Skudlarski, P., et al.: Brain connectivity is not only lower but different in schizophrenia: a combined anatomical and functional approach. *Biol. Psychiatry* 68(1), 61–69 (2010)
3. Tsiaras, V., et al.: Extracting biomarkers of autism from meg resting-state functional connectivity networks. *Comput. Biol. Med.* 41(12), 1166–1177 (2011)
4. Vissers, M., Cohen, M., Geurts, H.: Brain connectivity and high functioning autism: a promising path of research that needs refined models, methodological convergence, and stronger behavioral links. *Neurosci. Biobehav. Rev.* 36(1), 604–625 (2012)
5. Calhoun, V., Kiehl, K., Pearlson, G.: Modulation of temporally coherent brain networks estimated using ica at rest and during cognitive tasks. *Hum. Brain Mapp.* 29(7), 828–838 (2008)
6. Calhoun, V., Adali, T., Pearlson, G., Pekar, J.: A method for making group inferences from functional mri data using independent component analysis. *Hum. Brain Mapp.* 14(3), 140–151 (2001)
7. Yang, Z., Yuan, Z., Laaksonen, J.: Projective non-negative matrix factorization with applications to facial image processing. *Int. J. of Pattern Recog. and Artif. Intell.* 21(8), 1353–1362 (2007)
8. Batmanghelich, N., Taskar, B., Davatzikos, C.: Generative-discriminative basis learning for medical imaging. *IEEE Trans. Med. Imaging* 31(1), 51–69 (2011)
9. Montez, T., Linkenkaer-Hansen, K., van Dijk, B., Stam, C.: Synchronization likelihood with explicit time-frequency priors. *Neuroimage* 33(4), 1117–1125 (2006)
10. Yang, Z., Oja, E.: Linear and nonlinear projective nonnegative matrix factorization. *IEEE Trans. Neural Netw.* 21(5), 1734–1749 (2010)
11. Hyvarinen, A., Oja, E.: Independent component analysis: algorithms and applications. *Neural Netw.* 13(4-5), 411–430 (2000)