

Dynamic Connectivity Factorization: Interpretable decompositions of non-stationarity

Aapo Hyvärinen^{*†}, Junichiro Hirayama[†], and Motoaki Kawanabe[†]

^{*}Dept of Computer Science and HIIT, University of Helsinki, Finland

[†]Cognitive Mechanisms Laboratories, ATR, Kyoto, Japan

Abstract—In many multivariate time series, the correlation structure is non-stationary, i.e. it changes over time. Analysis of such non-stationarities is of particular interest in neuroimaging, in which it leads to investigation of the dynamics of connectivity. A fundamental approach for such analysis is to estimate connectivities separately in short time windows, and use existing machine learning methods, such as principal component analysis (PCA), to summarize or visualize the changes in connectivity. Here, we use the PCA approach by Leonardi et al as the starting point and present two new methods. Our goal is to simplify interpretation of the results by finding components in the original data space instead of the connectivity space. First, we show how to further analyse the principal components of connectivity matrices by a tailor-made two-rank matrix approximation, in which the eigenvectors of the conventional low-rank approximation are transformed. Second, we show how to incorporate the two-rank constraint in the estimation of PCA itself to improve the results. We further provide an interpretation of the method in terms of estimation of a probabilistic generative model related to blind source separation methods and ICA. Preliminary experiments on magnetoencephalographic data reveal possibly meaningful non-stationarity patterns in power-to-power coherence of rhythmic sources (i.e. correlation of amplitudes).

I. INTRODUCTION

Analysing non-stationarities of covariances, correlations, or other kinds of statistical connectivities is a topic of great interest both in machine learning and neuroimaging [1], [2], [4]. In the machine learning literature, emphasis has been placed on accurately estimating connectivities in a nonstationary environment [3]. In many applications, however, it is important that we can summarize or visualize the non-stationarity of the system in an intuitively appealing way. A simple way of summarizing non-stationary behaviour would be to compute some kind of connectivity statistics for different time segments, and then perform PCA in the space of those connectivity statistics [2]. The connectivity statistics could consist of covariance matrices or correlation matrices, for example.

However, computing principal components of connectivity matrices does not in itself lead to a very easily interpretable results, since the principal components are of the same form as the connectivity matrices. Connectivity matrices can be very complex and high-dimensional objects, and their visualization and analysis is a complex topic in itself.

Here, we develop a method for analysing the nonstationarity of connectivities based on components that are linear in the original data space. That is, we attempt to find pairs

of components of the original data which have maximally nonstationary behaviour, i.e. their connectivities are changing as strongly as possible. Such pairs are easy to interpret and to visualize in most cases.

II. BACKGROUND AND MOTIVATION

Denote by $C_\tau, \tau = 1, \dots, k$ a number of connectivity (correlation/covariance) matrices obtained from time segments of a multivariate time series $\mathbf{x}(t)$. A fundamental approach for analysing such matrices is to perform PCA on the vectorized forms of the matrices, i.e. considering the vectorized form of C_τ an observation at time point t , and performing PCA in the usual way. Thus, we obtain a number of principal component matrices which have the same form as C_τ . Let us denote by \mathbf{K} such a principal component.

A major problem is how to further analyse or visualize the obtained \mathbf{K} . A low-rank approximation of \mathbf{K} could be used to analyse its structure. A conventional approach is to use a one-rank approximation of \mathbf{K} by the dominant eigenvector. Using a single eigenvector is mainly able to represent groups of closely connected variables (“cliques”), such that the correlations between those variable change together. Likewise, a two-rank approximation mainly represents two groups which have connections inside each group. Such an analysis is useful in many cases, but we don’t consider it here for several reasons: 1) The principles of such low-rank approximations are already well-known, being accomplished simply by the eigen-value decomposition of \mathbf{K} .

2) What would be very interesting in brain imaging, is to find two groups of variables, such that the connectivity *between* the two groups is changing strongly, as opposed to changes of connectivity inside a group or groups of variables.

3) We have found empirically that the \mathbf{K} in real data often cannot be meaningfully represented by a conventional low-rank approximation. \mathbf{K} often has an eigenspectrum characterized by a *plus-minus structure*, i.e. one large positive eigenvalue, and one large negative one. The intuitive and practical meaning of such an eigenstructure needs to be understood.

The goal of section III below is to explain the meaning of such an plus-minus eigenstructure. We will see that it is in fact closely connected to point 2 above, i.e. modelling changing connectivity between two groups of variables. On the other hand, it is clear that we should try to improve the method by incorporating such low-rank constraints in the optimization problem itself; this will be done in Section IV below.

III. GROUP INTERACTIONS AS ORTHOGONAL TWO-RANK APPROXIMATION

A. Definition as optimization problems

Assume we are given a matrix \mathbf{K} which has the same dimensions as a connectivity matrix, as a result of PCA of the connectivity matrices or some similar method. In this section, we consider how to develop a low-rank approximation of \mathbf{K} which is particularly suitable for neuroimaging.

Denote by \mathbf{w} and \mathbf{v} two vectors in the data space which define brain areas or something similar. As an important departure from conventional analysis, let us also assume that \mathbf{w} and \mathbf{v} are *orthogonal*. This is because we want to analyse connectivities *between* two different groups of variables (brain areas). In conventional one-rank approximation, we would take the outer product of \mathbf{w} with itself, and we would be analysing connectivity *inside* a group of variables (those corresponding to non-zero entries in \mathbf{w}); a conventional two-rank approximation $\mathbf{w}\mathbf{w}^T + \mathbf{v}\mathbf{v}^T$ simply analyzes connections inside two groups separately from each other.

Thus, we use the outer product of \mathbf{w} and \mathbf{v} to model a pattern of connectivity between two regions. Due to symmetry of the connectivity matrices, we further use a symmetrized version. This leads to an optimization problem in which we attempt to find \mathbf{w} and \mathbf{v} by minimizing

$$\min_{\mathbf{w}^T \mathbf{v} = 0} \|\mathbf{K} - (\mathbf{v}\mathbf{w}^T + \mathbf{w}\mathbf{v}^T)\|^2 \quad (1)$$

where the norm is the Frobenius norm. This is a rather unconventional low-rank approximation since it uses the outer products of two orthogonal vectors. We call it the *orthogonal two-rank approximation*.

Some interpretations of the outer products in terms of brain imaging are shown in Fig. 1. In the most basic case (a), we can think of \mathbf{w} and \mathbf{v} as zero-one indicator functions of two brain regions or other groups of variables. Then, the idea is that it is the connectivity between those two areas that changes. On the other hand, if one of the vectors, say \mathbf{v} , has negative values as well, what we are modelling is *switching* of connectivity, or increase in one connectivity pattern coupled with reduction in another connectivity pattern, see (b). This could mean that in some parts of the data, there is connectivity between the area defined by \mathbf{w} and the area defined by the positive values of \mathbf{v} , while in other parts of the data, the connectivity is between the areas defined by \mathbf{w} and the negative parts of \mathbf{v} . (However, the exact interpretation depends on the baseline of connectivities, e.g. whether the connectivities are all non-negative.)

The optimization problem in (1) is clearly closely related to the following simple problem, reminiscent of PCA:

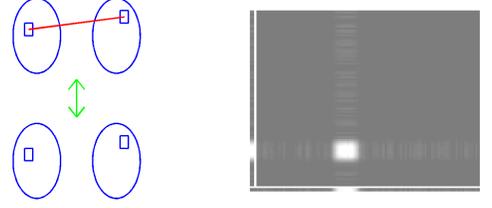
$$\max_{\|\mathbf{w}\|=\|\mathbf{v}\|=1, \mathbf{w}^T \mathbf{v}=0} \mathbf{w}^T \mathbf{K} \mathbf{v} \quad (2)$$

In the following we shall study these two simple optimization problems, and their connection.

B. Theory of orthogonal two-rank approximation

To solve the optimization problem in (2), we have the following theorem. (Proofs are omitted due to space constraints.)

a) Simple change of connectivity strength (\mathbf{w}, \mathbf{v} non-negative)



b) Switching of connectivity (\mathbf{w} non-negative, \mathbf{v} not)

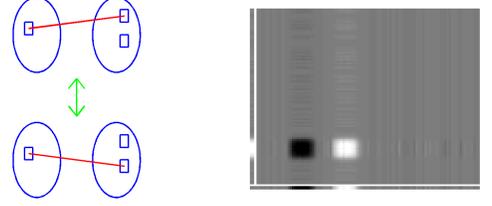


Fig. 1. Illustration of different connectivity pattern changes modelled by the orthogonal two-rank framework. Left: illustration of connectivities in terms of brain areas: each oval is one brain, boxes are brain areas, and red lines are connectivities. Right: the matrix produced as outer product of \mathbf{w}, \mathbf{v} . The vectors \mathbf{w} and \mathbf{v} are given as the line plots at the left and lower edges.

Theorem 1: Assume that \mathbf{K} is symmetric, and the largest and smallest eigenvalues of \mathbf{K} are distinct. (Here, largest and smallest means according to ordinary sorting, *not* using absolute values.) Then, the solution of the optimization problem (2) is given by

$$\mathbf{w} = \frac{1}{\sqrt{2}}(\mathbf{e}_{\max} + \mathbf{e}_{\min}), \quad \mathbf{v} = \frac{1}{\sqrt{2}}(\mathbf{e}_{\max} - \mathbf{e}_{\min}) \quad (3)$$

where \mathbf{e}_{\max} and \mathbf{e}_{\min} are the eigenvectors corresponding to the smallest and largest eigenvalues of \mathbf{K} . Denoting by λ_{\max} and λ_{\min} the largest and smallest eigenvalues of \mathbf{K} , the value of the objective at optimum is equal to $\frac{1}{2}(\lambda_{\max} - \lambda_{\min})$.

So, the solution of the optimization problem (2) is computationally simple, but quite surprising and unconventional. Next we consider the connection between the simple PCA-like objective in (2) and the orthogonal two-rank approximation.

Theorem 2: Consider the problem of approximating a symmetric matrix by a symmetrized outer product of two orthogonal vectors in (1).

- 1) The optimal \mathbf{w} and \mathbf{v} can be found, up to scaling constants, by solving the optimization problem (2) as in the Theorem 1.
- 2) The value of the objective at optimum is equal to $\|\mathbf{K}\|^2 - \frac{1}{2}(\lambda_{\max} - \lambda_{\min})^2$
- 3) The objective is 0 iff \mathbf{K} has rank two and $\lambda_{\min} = -\lambda_{\max}$.

The important implication of the theory presented above is that in our optimization problem, we do *not* use the eigenvectors themselves. This is in stark contrast to conventional low-rank approximations, which are based on using the outer

products of the eigenvectors with themselves. Further note that in Theorem 2, we cannot approximate every two-rank matrix exactly because we don't have separate scaling coefficients for the two eigenvectors. In a conventional two-rank approximation, we have the two eigenvalues as such scaling coefficients, and we would be able to exactly approximate any two-rank matrix. Here, we obtain an exact approximation only if the two eigen-values have opposite signs but equal absolute values. Thus, our approximation is exact only for matrices with the special plus-minus structure mentioned in Section II.

IV. CONSTRAINED PCA OBJECTIVE

The analysis above gave new insight into the matrix PCA results. Next, we propose to directly integrate the orthogonal two-rank assumption into a PCA objective function.

In one time segment, the connectivity between the two areas can be defined as $\mathbf{v}^T \mathbf{C}_\tau \mathbf{w}$ which, in the case of covariance matrices, is actually the covariance of $\mathbf{v}^T \mathbf{x}$ and $\mathbf{w}^T \mathbf{x}$ in the time segment. We want to maximize the variance of this, to find components which explain as much of the nonstationary connectivity structure as possible. We have to constrain \mathbf{w} and \mathbf{v} in some way, so let us set their norms to be equal to unity. Thus, we obtain the (preliminary form of the) optimization problem

$$\max_{\|\mathbf{w}\|=\|\mathbf{v}\|=1, \mathbf{w}^T \mathbf{v}=0} \sum_{\tau=1}^k (\mathbf{w}^T \mathbf{C}_\tau \mathbf{v})^2 - \left(\sum_{\tau} \mathbf{w}^T \mathbf{C}_\tau \mathbf{v} \right)^2 \quad (4)$$

Furthermore, let us assume that the mean connectivities have been subtracted to obtain

$$\tilde{\mathbf{C}}_\tau = \mathbf{C}_\tau - \frac{1}{k} \sum_{i=1}^k \mathbf{C}_i \quad (5)$$

so that $\sum_{\tau} \tilde{\mathbf{C}}_\tau = \mathbf{0}$. Thus, we obtain the final optimization problem as

$$\max_{\|\mathbf{w}\|=\|\mathbf{v}\|=1, \mathbf{w}^T \mathbf{v}=0} \sum_{\tau=1}^k (\mathbf{w}^T \tilde{\mathbf{C}}_\tau \mathbf{v})^2 \quad (6)$$

It is useful to consider the basic case of two connectivity matrices, i.e. $k = 2$. Then the subtraction of means means that $\tilde{\mathbf{C}}_1 = -\tilde{\mathbf{C}}_2$, and we simply have $\mathbf{K} = \tilde{\mathbf{C}}_1$ up to a scaling constant. In fact, the constrained PCA is then simply based on analysing the difference of the original connectivities since $\mathbf{K} = \frac{1}{2}(\mathbf{C}_1 - \mathbf{C}_2)$, and we obtain directly the optimization problem in (2). However, with $k \geq 3$, we need to develop a new algorithm.

It is reasonable to start the optimization at a point given by the orthogonal two-rank approximation of matrix PCA, which means that maximization of this objective can be considered a *finetuning* of matrix PCA results. We have developed a dedicated alternating variables algorithm for solving this optimization problem, in which both alternating steps are in closed-form and no gradient steps are needed, but a description is omitted due to lack of space.

The objective in (6) above considers a single pair of components of connectivities. Further pairs can be obtained by

well-known deflationary methods, i.e. optimizing the objective under the constraint that the new vectors are orthogonal to those already estimated.

V. PROBABILISTIC GENERATIVE MODEL

To get further insight into our methods, we have developed a generative probabilistic model such that the methods above can be considered to estimate parameters in it. Due to space constraints, we don't go into details but just mention the basic idea here.

Assume we have two time segments and in both, the data follows a linear mixing model, with standardized, possibly Gaussian components s_i

$$\mathbf{x} = \mathbf{A} \mathbf{s} \quad (7)$$

where the number of s_i is equal to the number of the x_i , and \mathbf{A} is square. Assume the matrix \mathbf{A} is the same for the whole data. Assume that in the first segment, the s_i are independent. To model changes in connectivity, assume that in the second segment, we have the perfect correlation $s_1 = s_2$ but otherwise the components are independent. Then we have

$$\mathbf{C}_1 = \mathbf{A} \mathbf{A}^T, \quad \mathbf{C}_2 = \mathbf{A} \mathbf{A}^T + \mathbf{a}_1 \mathbf{a}_2^T + \mathbf{a}_2 \mathbf{a}_1^T \quad (8)$$

$$\mathbf{K} = \mathbf{C}_2 - \mathbf{C}_1 = \mathbf{a}_1 \mathbf{a}_2^T + \mathbf{a}_2 \mathbf{a}_1^T \quad (9)$$

so, by Theorem 2, optimization of (2) will find $\mathbf{w} = \mathbf{a}_1$ and $\mathbf{v} = \mathbf{a}_2$. This shows that the methods considered in this paper can be seen to analyse changes in correlation structure of components in ICA-like models. Although we assumed above that the components are conditionally Gaussian in each segment, the global distribution of the components is non-Gaussian due to being a Gaussian scale mixture.

VI. EXPERIMENTS ON MAGNETOENCEPHALOGRAPHIC DATA

To validate the methods on real data, we used magnetoencephalographic (MEG) data from [6]. The data comprised a single 12 minute session recorded on a Elekta Neuromag 306-channel neuromagnetometer. The subject received alternating stimulation of visual, auditory, and tactile modalities, interspersed with rest periods.

We first Morlet-filtered the data with a center-frequency of 10 Hz to extract alpha-range activity. We next performed ICA on the Morlet-filtered gradiometers to separate sources of rhythmic activity. Finally, we computed the amplitudes (envelopes) of the independent components, and used these as input to the method proposed here.

Correlation coefficient matrices were computed in non-overlapping time windows of a length of 5 seconds. The two variants of the method, 1) Matrix-space PCA followed by two-rank approximation (Section III), and 2) Constrained PCA (Section IV) were applied on the data.

We show the first component pair (\mathbf{w}, \mathbf{v}) for the two methods in Fig. 2, and the second component pair in Fig. 3. The components are visualized on the measurement helmet by adding together the spatial patterns (squares of columns of mixing matrix) of those underlying independent components

1st component pair for MEG amplitudes

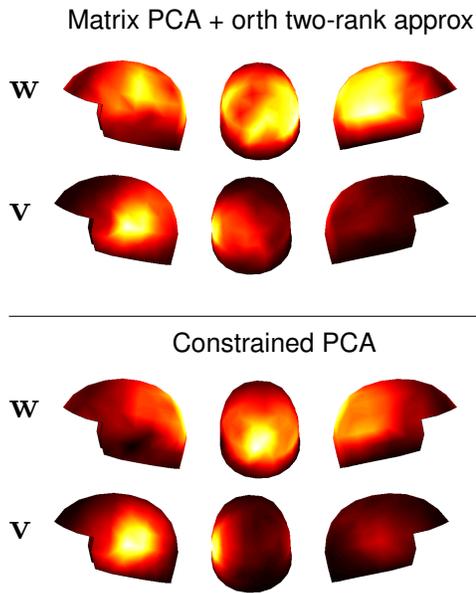


Fig. 2. Results on MEG amplitudes: the first principal component pair obtained by the two methods. Top panel: Matrix principal component analysed by orthogonal two-rank approximation. Bottom panel: Constrained PCA results. In each panel, the upper and lower parts give the two components in the pair obtained.

that contribute to each connectivity component. Thus we obtain a rough plot of the spatial extents of the components. Note that w and v could be interchanged, and we have here manually switched them in some plots to make the results of the two methods as similar as possible, purely for the ease of visualization.

In general, the results using constrained PCA look cleaner and more plausible. The first component pair (Fig. 2, lower panel), shows that the strongest non-stationarity is in the connectivity between an occipito-parietal area and the left temporal area. The second component pair (Fig. 3, lower panel) shows that another strongly nonstationary connectivity is between inferior occipital areas and slightly superior occipito-parietal areas.

VII. DISCUSSION

We proposed a method, Dynamic Connectivity Factorization, for analysing the dynamics of connectivity patterns in terms of components which are linear in the original data. The goal is to analyse the nonstationarities in a way which is intuitively comprehensible and easy to visualize. To this end, we find two linear, orthogonal components of the data such that the connectivity between them is maximally non-stationary. The orthogonality constraint sets the method apart from related methods, such as common spatial patterns (CSP), or blind source separation using non-stationarity [5].

2nd component pair for MEG amplitudes

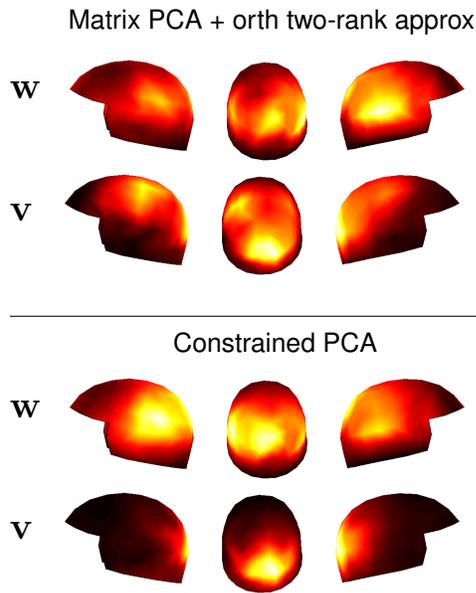


Fig. 3. Results on MEG amplitudes: the second principal component pair obtained by the two methods. (See Fig. 2 for legend.)

We presented two variants of the method, one based on further analysing the PCA of connectivity matrices [2] by a dedicated, orthogonal two-rank approximation, and another based on formulating a constrained PCA objective function and optimizing it by a tailor-made algorithm. The latter seemed to produce better results on MEG, although it has to be noted that the experiments were quite preliminary and the validation of the results purely visual. Although we only presented MEG results here, we hope the method will be applicable to fMRI and other brain imaging modalities.

REFERENCES

- [1] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun. Tracking whole-brain connectivity dynamics in the resting state. *Cerebral Cortex*, page bhs352, 2012.
- [2] N. Leonardi, J. Richiardi, M. Gschwind, S. Simioni, J.-M. Annoni, M. Schluep, P. Vuilleumier, and D. Van De Ville. Principal components of functional connectivity: A new approach to study dynamic brain connectivity during rest. *NeuroImage*, 83:937–950, 2013.
- [3] S. Liu, J. A. Quinn, M. U. Gutmann, and M. Sugiyama. Direct learning of sparse changes in markov networks by density ratio estimation. In *Machine Learning and Knowledge Discovery in Databases (ECML2013)*, pages 596–611. Springer, 2013.
- [4] X. Liu and J.H. Duyn. Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proceedings of the National Academy of Sciences*, 110(11):4392–4397, 2013.
- [5] K. Matsuoka, M. Ohya, and M. Kawamoto. A neural net for blind separation of nonstationary signals. *Neural Networks*, 8(3):411–419, 1995.
- [6] P. Ramkumar, L. Parkkonen, R. Hari, and A. Hyvärinen. Characterization of neuromagnetic brain rhythms over time scales of minutes using spatial independent component analysis. *Human Brain Mapping*, 33(7):1648–1662, 2012.