Journal of Neuroscience Methods 271 (2016) 92–106

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

p-Value combiners for graphical modelling of EEG data in the frequency domain☆

Deborah Schneider-Luftman

Department of Mathematics, 180 Queen's Gate, Imperial College London, SW7 2BB London, UK

HIGHLIGHTS

- We adapt *p*-value combiners to work for frequency domain data.
- A two-step procedure for graphical modelling on EEG data is proposed.
- We control for false detections across networks with various level adjustments.
- Methods are proposed to combine results across subjects and create group results.
- Coloured graphs are made for each group, showing the prevalence of connections.

ARTICLE INFO

Article history:
Received 22 February 2016
Received in revised form 24 June 2016
Accepted 18 July 2016
Available online 21 July 2016

Keywords:
Partial coherence
Multivariate time series
Multiple hypothesis testing
p-Value combiners
Electroencephalogram (EEG) data

ABSTRACT

Background: In the graphical modelling of brain data, we are interested in estimating connectivity between various regions of interest, and evaluating statistical significance in order to derive a network model. This process involves aggregating results across frequency ranges and several patients, in order to obtain an overall result that can serve to construct a graph.
New method: In this paper, we propose a method based on *p*-value combiners, which have never been used in applications to EEG data analysis. This new method is split into two aspects: frequency-wide tests and group-wide tests. The first step can be effectively adjusted to control for false detection rate.
Results: This two-step protocol is applied to EEG data collected from distinct groups of mental health patients, in order to draw graphical models for each group and highlight structural connectivity differences. Using the method proposed, we show that it is possible to reliably achieve this while effectively controlling for false connections detection.
Comparison with existing method(s): Conventionally, the Holm's Stepdown procedure is used for this type of problem, as it is robust to type I errors. However, it is known to be conservative and prone to false negatives. Furthermore, unlike the proposed methods, it does not directly output a decision rule on whether to accept or reject a statement.
Conclusions: The proposed methodology offers significant improvements over the stepdown procedure in terms of error rate and false negative rate across the network models, as well as in term of applicability.

1. Introduction

The understanding of connectivity in large-dimensional time series has been a topic of central importance in neurology, and more precisely in neurological imaging. The interest in these techniques is widespread across imaging techniques (EEG, Medkour et al., 2010; fMRI, Marrelec et al., 2006) and experimental works of various types (learning experiments, Fiecas and Ombao, 2014; motor skills, Mima et al., 2000; resting-state, Salvador et al., 2005). One of the most important features of neurological data analysis is functional connectivity between parts of the brain: how do various regions of interest interact? For this purpose, graphical modelling of time series is an ideal tool.

A well-known contribution to this field is the frequency-domain approach exposed in Dahlhaus (2000). In this methodology, the data is transformed into the frequency domain, where its co-dependency structure is analysed via the partial coherences. The partial coherence measures the connection between two series after the removal of the linear effects of the remaining series. It is a function of frequency and can be used to reflect connectivity across any frequency range $\Omega \subseteq [0, f_N]$. It is derived from $\mathbf{S}^{-1}(f)$, the inverse of the spectral matrix $\mathbf{S}(f)$. Its use in network analysis and

☆ This work is supported by EPSRC (UK) via a PhD grant.
E-mail address: ds711@ic.ac.uk

http://dx.doi.org/10.1016/j.jneumeth.2016.07.006

related methodologies has become very widespread in the literature (Makhtar et al., 2014; Mima et al., 2000; Pohja et al., 2002).

Once the partial coherences have been estimated, an important question arises concerning the statistical significance of the results. A zero-valued estimate would indicate that no connection is detectable between two variables at a given frequency. The standard approach is to test for the following at all frequencies f in range:

H_0 : partial coherence $= 0$ vs H_1 : partial coherence > 0.

In some cases, multiple measurements of the data are available, hence allowing the estimation of a bootstrapped distribution (Fiecas and Ombao, 2011). When this is not possible, the use of analytical tools is required. A popular alternative involves the Partial Mutual Information (PMI), which integrates the estimates at all frequencies within a band into one variable for each partial coherence, and is then compared to a threshold (Salvador et al., 2005). However, its distribution and statistical properties remain largely unknown, and it is often unclear how the threshold value should be set for a given band in $[0, f_N]$.

In this paper, we consider an alternative protocol that follows the method detailed in Medkour et al. (2009). Under conventional pre-processing and estimation methods for the spectral matrix $\mathbf{S}(f)$, its distribution and the distribution of its partial coherence are known (Goodman, 1963), and then, under H_0, the partial coherence estimates can be modelled as Beta-distributed random variables.

In the frequency domain approach, results are derived at each frequency f, for all frequencies in range $f \in \Omega$. However, this is many steps away from an overall graphical model. Once the partial coherences have been measured and tested for significance at a suitable level α, how can the results be aggregated across frequencies and subjects to deliver one graph?

This can be regarded as a multiple hypothesis testing problem, otherwise known as a conjunction analysis problem in Neurology (Friston et al., 1999). A traditional approach to this is the Holm's Stepdown procedure (Lehmann and Romano, 2005; Holm, 1979), where the null hypothesis H_0 is tested for at each frequency ordered by p-values $p_{(1)} \leq p_{(2)} \leq \cdots$. Every time H_0 is rejected, the procedure moves on to the next frequency, and only stops at the frequency L where H_0 is finally accepted. This method is robust to Type I errors, as it is designed to control the family-wise error rate (FWER) below a desired level α (Lehmann and Romano, 2005). However, the stepdown is prone to false negatives. Furthermore, translating the resulting L into a decision is ambiguous. L may sometimes be very small, especially compared to the number of frequencies it is computed for. For instance, if $L = 1$ for some pairs of variables, should the connections between these pairs still be included in the graphical model? No substantial research has been carried out to answer this beyond some case-specific solutions (Medkour et al., 2010).

Multiple hypothesis testing is not limited to the use of the Stepdown procedure. In other applications, the use of p-value combiners is very prevalent. In Genomics, the Westfall–Young min-p procedure (Westfall and Young, 1993) proves to be very popular, as it is robust to Type I errors and can also handle correlated data by estimating the joint H_0 distribution through resampling. Other well-known combiners rely on Bayesian inference, such as Efron's empirical Bayes method (Efron, 2003), where prior probabilities are assigned on the proportion of null and non-null statements and the false discovery rate is evaluated empirically.

In the context of spectral domain analysis, closed-form analytical methods tend to be preferred, due to the large computational cost associated with performing calculations at each frequency. In this category of methods, The Fisher (Fisher, 1932) and Simes (Simes, 1986) combiners constitute popular examples, that are widely used in applications of Computational Statistics, such as Genetic Epidemiology (Sungho et al., 2009) and Biostatistics (Chen et al., 2014). They deliver a single scalar that can then be tested on well-defined distributions, in order to ascertain the significance of an overall proposition. The use of these p-value combiners has been relatively rare in graphical modelling of neurological data thus far. Conventionally, they require that the set of multiple tests are independent, which is almost never the case with frequency domain data. However it is possible to generalise their use for this specific application.

In this paper, we review various p-values combiners and assess their suitability for graphical modelling of EEG data compared to the Stepdown procedure. After reviewing some background results in frequency domain analysis and multiple hypothesis testing, we propose a two-step procedure to carry out graphical modelling on EEG data.

- We demonstrate how the classical p-value combiners can be used on a subset of the frequency range that only includes uncorrelated data, and evidence their performance on simulated data.
- Test combiners can also be used to ascertain the significance of a graphical model for a sample population. Using EEG measurements from three distinct groups of mental health patients, we demonstrate how we can aggregate the results of each patient in all groups in order to obtain a group-wide coloured graphs, which show the intensity of connections in each group.
- When combining results across a frequency range, each connection between a pair of channel is evaluated independently. In doing so, it is important to control for the detection of false positives when constructing individual graphs. We show how this can be managed using a false edge detection adjustment, for both low-dimensional and larger dimensional data.

2. Background – graphical modelling

Let $\{\mathbf{X}_t\}$ be a 2nd order stationary vector time series, $\{\mathbf{X}_t\} \in \mathbb{R}^p$, $t \in \{0, \ldots, T-1\}$, with an associated spectral matrix $\mathbf{S}(f) \in \mathbb{C}^{p\times p}$. Many estimation procedure exist for $\mathbf{S}(f)$, here we choose the multitaper spectral estimate for its good statistical and analytical properties (Percival and Walden, 1993). It starts with a set of K orthogonal tapers, satisfying the following property:

$$\sum_{t}^{T} h_{t,k}h_{t,l} = \begin{cases} 0 & \text{forall } k \neq l, \\ 1 & \text{if } k = l. \end{cases}$$

There are many types orthogonal tapers $\{h_{t,k}\}$ that are regularly used in the literature on spectral estimation, we choose here the Sine tapers, for their ease of implementation (Walden et al., 1995), defined for all $k \leq K$ and $t \leq T-1$:

$$h_{t,k} = \sqrt{\frac{2}{T+1}} \sin\left(\frac{(k+1)\pi t}{T+1}\right), \qquad t \leq T, \quad k \leq K$$

Using the multiple tapers $\{h_{t,k}\}$, we can define the following Fourier transforms $\mathbf{J}_k(f)$ for $k < K$ on the data $\mathbf{X}_t$:

$$\mathbf{J}_k(f) = \Delta_t \sum_{t} h_{k,t}\mathbf{X}_t e^{-\mathbf{i}2\pi f t\Delta_t} \in \mathbb{C}^{p\times 1},$$

which in turn can be used to create an estimate for the spectral matrix $\mathbf{S}(f)$, called the multitaper estimate:

$$\hat{\mathbf{S}}(f) = \hat{\mathbf{S}}_K^{(mt)}(f) = \frac{1}{K}\sum_{k=1}^{K} \mathbf{J}_k(f)\mathbf{J}_k^H(f). \tag{2.1}$$

To ensure the invertibility of the matrix $\mathbf{S}(f)$, we require that the number of tapers exceed the dimensions of the data, i.e. $K > p$.