ORIGINAL PAPER

On the "Dependence" of "Independent" Group EEG Sources; an EEG Study on Two Large Databases

Marco Congedo · Roy E. John · Dirk De Ridder ·
Leslie Prichep · Robert Isenhart

Abstract The aim of this work is to study the coherence profile (dependence) of robust eyes-closed resting EEG sources isolated by group blind source separation (gBSS). We employ a test–retest strategy using two large sample normative databases (N = 57 and 84). Using a BSS method in the complex Fourier domain, we show that we can rigourously study the out-of-phase dependence of the extracted components, albeit they are extracted so as to be in-phase independent (by BSS definition). Our focus on lagged communication between components effectively yields dependence measures unbiased by volume conduction effects, which is a major concern about the validity of any dependence measures issued by EEG measurements. We are able to show the organization of the extracted components in two networks. Within each network components oscillate coherently with multiple-frequency dynamics, whereas between networks they exchange information at non-random multiple time-lag rates.

Introduction

The functional organization of the brain at rest is currently conceived in terms of resting state networks (RSNs), clusters of brain regions, mostly cortical, inter-connected anatomically and functionally (Beckmann et al. 2005; Damoiseaux et al. 2006; Fox and Raichle 2007; Mantini et al. 2007; van den Heuvel et al. 2008). The study of RSNs has shifted the focus in neuroimaging from the exact localization of specialized brain functions (looking for "things in a place") to the understanding of the interplay of widespread brain structures.

Regarding the study of RSNs by EEG, biophysical and neurophysiological studies suggest that each RSN may exhibit complex dynamic associated with multiple frequencies simultaneously (Mantini et al. 2007). Studying the distribution of scalp EEG power at rest, as in Chen et al. (2008), does not allow the study of RSNs by EEG because scalp voltage is a mixing of underlying source activity (volume conduction: see Nunez and Srinivasan 2006) and because scalp power is not a comprehensive measure of widespread coherent synchronization. Instead, it appears more appropriate extracting robust synchronizations all over the cortex and testing them altogether (the whole cohort) along the frequency dimension. The multivariate tool permitting such investigation in large samples of individuals is group independent component analysis (gICA) (Calhoun et al. 2001; Schmithorst and Holland 2004). Whereas ICA relies on high-order statistics, here we use another blind source separation (BSS) framework

M. Congedo (✉)
Gipsa-lab, National Center for Scientific Reserach (cnrs),
University Joseph Fourier, University Stendhal,
Grenoble Institute of Tecnology, 961 rue de la Houille Blanche,
Domaine Universitaire, BP 46-38402, Grenoble, France
e-mail: marco.congedo@gmail.com

R. E. John · L. Prichep · R. Isenhart
Department of Psychiatry, Brain Research Laboratory,
New York University Medical School, New York, USA

D. De Ridder
Brain Research center Antwerp for Innovative and
Interdisciplinary Neuromodulation (BRAI²N) and Department
of Neurosurgery, University and Hospital of Antwerp,
Edegem, Belgium

based on second-order statistics. We have previously argued that second-order BSS methods suits well EEG data (Congedo et al. 2008). Importantly, it allows us to address the problem of possible dependence among the extracted regions of coherent synchronization.

Group BSS extracts scalp spatial maps and associated EEG time-courses, referred together to as components. We wish to extract components that can be found on a substantial proportion of normal individuals at rest. Since there is no way to establish a-priori how many of such components can be established, nor if they are reliable, we employ a test–retest strategy using two independent large sample normative databases (N = 57 and 84) and retain as many components as we can replicate, proceeding in descending order of explained variance; a similar strategy has been previously employed (Damoiseaux et al. 2006) in an fMRI study. Once robust normative components are extracted, we characterize the cortical structures involved in each component by a distributed source localization of the spatial maps (Greenblatt et al. 2005) and their spectral profile. We are able to replicate on the two databases seven components with nearly identical spatial and frequency distribution (detailed in a separate paper). In this paper we study the out-of-phase (lagged) coherence (dependence) of the extracted components using recent advances on connectivity measures adapted to EEG data. We are able to show the organization of the components in two networks. Within each network components oscillate coherently with multiple-frequency dynamics, whereas between networks they exchange information at non-random multiple time-lag rates.

Materials and methods

It is well known that spectral EEG measures follow developmental equations, meaning that the frequency composition of the EEG reflects the maturational status of the brain (John et al. 1980). In order to avoid confounding age effects we consider in this study only adult individuals between 17 and 30 years old. Two independent normative databases previously acquired were used for this study. One is a subset of the normative database of the Brain Research Laboratory (BRL), New York University (N = 57; age range 17–30) and the other the adult normative database of Nova Tech EEG (NTE), Inc., Mesa, AZ (N = 84; age range 18–30). Exclusion criteria for the BRL database were known psychiatric or neurological illness, drug/alcohol abuse, current psychotropic/CNS active medications, history of head injury (with loss of consciousness) or seizures. Exclusion criteria for the NTE database were a psychiatric history in any relative and participant of drug/alcohol abuse, head injury (at any

age, even very mild), headache, physical disability and epilepsy.

Recording procedures and settings were very similar for the two databases. In both cases about 3–5 min of EEG was continuously recorded while participant sat with the eye-closed on a comfortable chair in a quiet and dimly lit room. EEG data were acquired at the 19 standard leads prescribed by the 10–20 international system (FP1, FP2, F7, F3, FZ, F4, F8, T3, C3, CZ, C4, T4, T5, P3, PZ, P4, T6, O1, O2) using both earlobes as reference and enabling a 60 Hz notch filter to suppress power line contamination. The impedance of all electrodes was kept below 5 K Ω. Data of the NTE database were acquired using the 12-bit A/D NeuroSearch-24 acquisition system (Lexicor Medical technology, Inc., Boulder, CO) and sampled at 128 Hz, whereas data of the BRL database were acquired using the 12-bit A/D BSA acquisition system (Neurometrics, Inc., New York, NY) and sampled at 100 Hz. For consistency, we subsequently up-sampled the BRL database to 128 Hz using a natural cubic spline interpolation routine (Congedo et al. 2002). In order to minimize inter-subject variability we removed from all data any biological, instrumental and environmental artifacts, paying particular attention to biological artifacts generated by the eyes, the hearth and the muscles of the neck, face and jaw. EEG recordings were visually inspected on a high-resolution screen and epochs containing visible artifacts were marked and ignored for ensuing analysis.

Frequency-Domain Statistics

All statistics used in this study are summarized in the complex hermitian Fourier cross-spectral matrices $\mathbf{S}_f \in \mathbb{C}^{\mathbb{E} \cdot \mathbb{E}}$, where f is the discrete frequency index and E the number of electrodes (Bloomfield 2000). We can write $\mathbf{S}_f = \mathbf{C}_f + i\mathbf{Q}_f$, where $i = \sqrt{-1}$. Symmetric $\mathbf{C}_f \in \Re^{E \cdot E}$, referred to as the *cospectral matrix*, holds in the main diagonal the power spectra and in the off-diagonal elements the in-phase (or with a half cycle phase shift, i.e., opposite sign) dependency structure. Antisymmetric $\mathbf{Q}_f \in \Re^{E \cdot E}$, referred to as the *quadrature spectral matrix*, holds in the off-diagonal elements the out-of-phase (a quarter cycle in either direction) dependency structure. Both dependency structures are of second order. The cospectral matrix is equivalent to the covariance matrix of the data band-pass filtered for its discrete frequency: by Parseval's theorem the sum of all co-spectral matrices is equivalent to the data covariance matrix. We estimate individual cross-spectral matrices as the average obtained by Fast Fourier Transform (FFT) on 50% sliding overlapping 2 s windows tapered using the function introduced by Welch (1967). Finally, we obtain group average cospectral matrices as the grand average across individuals.

Group Blind Source Separation

For E scalp sensors and $M \leq E$ EEG dipolar fields to be estimated, the BSS linear model employed describes the superposition principle, i.e., we state

$$\mathbf{v}(t) = \mathbf{A}\mathbf{s}(t), \tag{1}$$

where $\mathbf{v}(t) \in \Re^E$ is the sensor measurement vector, $\mathbf{A} \in \Re^{E \cdot M}$ is a time-invariant non-singular mixing matrix and $\mathbf{s}(t) \in \Re^M$ holds the time-course of the source components. Note that model 1 describes the instantaneous (in-phase) diffusion of current source over measurement sites, in fact describing the effect of direct current and volume conduction (Congedo et al. 2008). Our source estimation is given by

$$\hat{\mathbf{s}}(t) = \mathbf{B}\mathbf{v}(t), \tag{2}$$

where separating matrix $\mathbf{B} \in \Re^{M \cdot E}$ is what we wish to estimate by BSS assuming no knowledge of A and weak assumptions on s.

A wide array of BSS/ICA methods exist. In this work we use a method based on the approximate joint diagonalization (AJD) of Fourier cospectral matrices, which is a robust and computationally fast approach (Congedo et al. 2008). We diagonalize grand-average cospectral matrices in the frequency range 0.5–30 Hz only, which is the range with highest signal-to-noise ratio. This results in $F = 60$ frequencies with 0.5 Hz resolution (0.5, 1, 1.5, …, 30 Hz). This approach is analogous to the averaging group ICA approach described for fMRI by Schmithorst and Holland (2004). The weak assumption on the source process we need to solve the BSS problem with this approach is that sources are uncorrelated and have non-proportional power spectrum (Congedo et al. 2008).

In order to estimate $M < E$ source components the matrix B is found with a classical two-stage process, which allows the estimation of the M most energetic components while reducing the noise. This involves first whitening the grand-average cospectral structure and then finding a matrix B best diagonalizing (AJD) all reduced cospectra. To solve the AJD problem we use an iterative algorithm previously developed (Pham and Congedo 2009), which is computationally fast. The reader is referred to (Congedo et al. 2008) for further details on AJDC.

Lagged Coherence

We estimate the lagged (out-of-phase) coherence between all 21 pairs of the seven components for all frequencies in between 0.5 and 55 Hz. To obtain source coherence let $c_f(x)$ be the xth diagonal element of $\mathbf{B}\mathbf{C}_f\mathbf{B}^T$, $c_f(xy)$ the off-diagonal elements of $\mathbf{B}\mathbf{C}_f\mathbf{B}^T$ and $q_f(xy)$ the off-diagonal

elements of $\mathbf{B}\mathbf{Q}_f\mathbf{B}^T$. The well-known squared coherence at discrete frequency f is

$$\rho_f^2(xy) = c_f^2(xy) + q_f^2(xy) \Big/ c_f(x)c_f(y), \tag{3}$$

Such measure can be conceived as the frequency-domain analogous of the coefficient of determination, the square of the Pearson product-moment correlation coefficient, but describes both in-phase *and* out-of-phase correlation. The measure is comprised between 0 (no correlation at discrete frequency f) and 1 (full correlation at discrete frequency f). It has been intensely used in EEG literature to describe the dependency between two scalp locations, however it is inflated by the artificial in-phase (zero-lag) correlation engendered by volume conduction (Nunez and Srinivasan 2006). Recently Nolte et al. (2004) proposed to consider instead the "imaginary part" of the coherency (the non-squared coherence)

$$\varphi_f(xy) = q_f(xy) \Big/ \sqrt{c_f(x)c_f(y)}, \tag{4}$$

which is not affected by volume conduction since it reflects only lagged correlation. Finally, Pascual-Marqui (2007) showed that the standard coherence 3 can be split in two additive parts: an instantaneous (in-phase) and a lagged (out-of-phase) part. The correct lagged coherence is given by

$$\phi_f^2(xy) = q_f^2(xy) \Big/ \Big[c_f(x)c_f(y) - c_f^2(xy) \Big]. \tag{5}$$

Such measure is particularly adapted in this context since our BSS method assumes null zero-lag correlation among sources. When applied in the source space, as we are doing, lagged coherence can be interpreted as the amount of cross-talk between the regions contributing to the source activity. Since the two components oscillate coherently with a lag, the cross-talk can be interpreted as information sharing by axonal transmission. The discrete Fourier transform decomposes the signal in a finite series of cosine and sine waves (in-phase and out-of-phase carrier waves, respectively, forming the real and imaginary part of the Fourier decomposition) at the so-called Fourier frequencies (Bloomfield 2000). The lag of the cosine waves with respect to their sine counterparts is inversely proportional to their frequency and amounts to a quarter of the period; for example, the period of a sinusoidal wave at 10 Hz is 100 ms. The sine is shifted a quarter of a cycle (25 ms) with the respect to the cosine. Then the lagged coherence 5 at 10 Hz indicates coherent oscillations with a 25 ms delay, while at 20 Hz the delay is 12.5 ms, etc.

The Φ_f^2 threshold of significance for a given α value is given by (Pascual-Marqui 2007)

$$th^{\phi_f^2}(\alpha) = 1 - e^{\left(-F^{-1}\left(\chi_{(1)}^2 (1-\alpha) \right) \Big/ N \right)}, \tag{6}$$

where $F^{-1}\left(\chi^2_{(1)}(1-\alpha)\right)$ is the inverse CDF of a Chi-square random variable with one degree of freedom.

Results

Figure 1 (top and center) shows the estimated lagged squared coherence 5 obtained on the two databases for all component pairs along frequencies, with significant results at the 0.2 level plotted using thick lines. The profiles appear clearly non-random and seem to concentrate in discrete frequency regions of high-communication rate, interleaved by regions of low-communication rate. The coherence profile displays peaks occurring at nearly identical frequencies in the two databases, although with inconsistent relative amplitude. The communication protocol among components displays a dynamic functioning with most component pairs interplaying at multiple frequencies (multiple time-lags) simultaneously, although most pairs prefer a narrow-band communication window. Almost all peaks found in the BRL databases can be found in the NTE database and viceversa. Figure 1 (bottom) shows the average between the two databases. Although no smoothing along adjacent frequencies has been performed, due to the size of our sample the coherence profiles show an acceptable signal-to-noise ratio. For the combined coherence (N = 141) we computed asymptotic significant values for type II error (α) equal to 0.2, 0.1 and 0.05. Significant coherences are reported in Fig. 2 in the form of a connectivity graph. We see that the seven components are organized in two independent (both in-phase and out-of-phase) networks, whereas significant out-of-phase cross-talk exists within each network.

Discussion

In this study we have shown that

1) an adequate BSS method not only allows canceling the volume conduction effect, i.e., the isolation of the relevant brain sources generating the observed scalp EEG, but also the study of residual (out-of-lag) cross-talk between such sources, unbiased by volume conduction.
2) group BSS can be purposefully used in EEG as group ICA is used in fMRI; we can replicate a set of seven components with nearly identical spatial and spectral content (data not shown). Furthermore, the organization of these components in networks communicating by specific frequency and time-lag protocols is also replicable.

In our two databases the significant cross-talk appears at low frequencies (≤ 11 Hz), which is in line with the role of
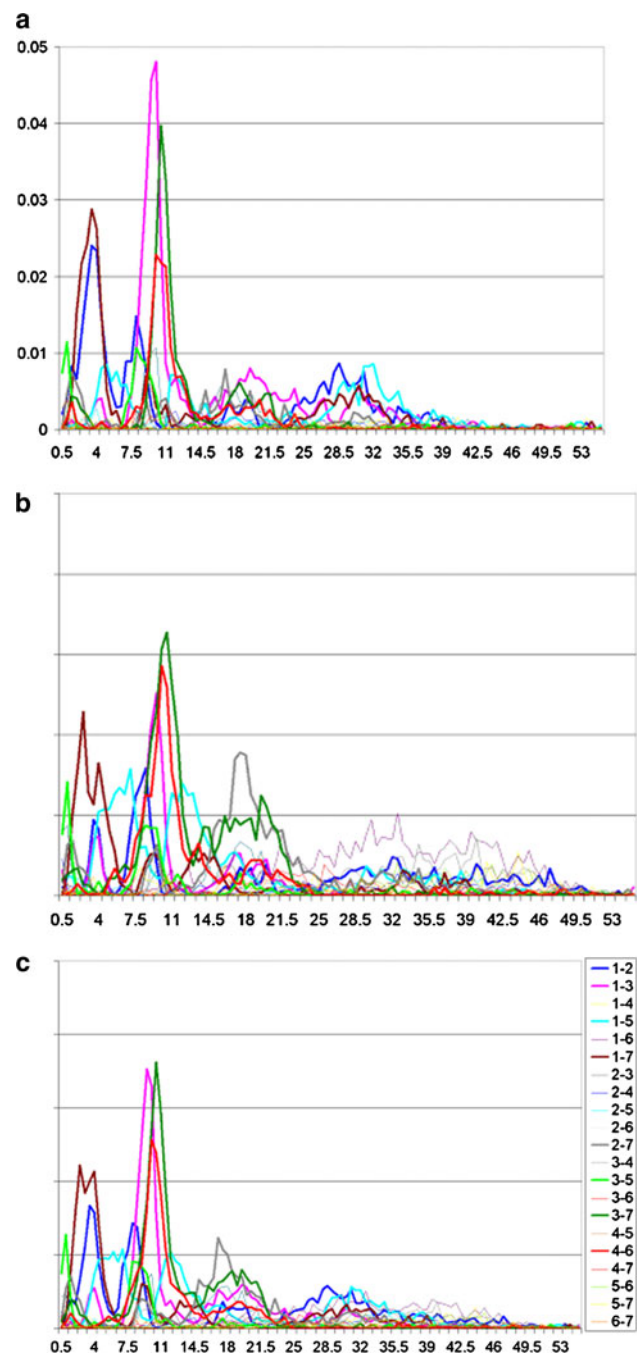
Fig. 1 Lagged squared coherence (5) in the 0.5–55 Hz range (0.5 Hz resolution) for all pair-wise couples of the seven components for the BRL database (*top*), NTE database (*center*) and the combination of the two (*bottom*). The coherence profiles of significant ($P < 0.2$) component pairs are drawn using a *thick line*

such frequencies in the long-(spatial) range co-ordination (Broyd et al. 2008; Fox and Raichle 2007). In fact both networks solicit simultaneously anterior and posterior regions and more in general regions far apart several centimeters from each other (data not shown), corroborating the hypothesis that axonal transmission is at the origin of the observed interplay.

Fig. 2 Significant lagged coherence (5) for three type II error rates (α). Components with at least two significant connections at the 0.1 level and at least three significant connections at the 0.2 level are marked as "hubs". The seven components organize in two fully independent networks
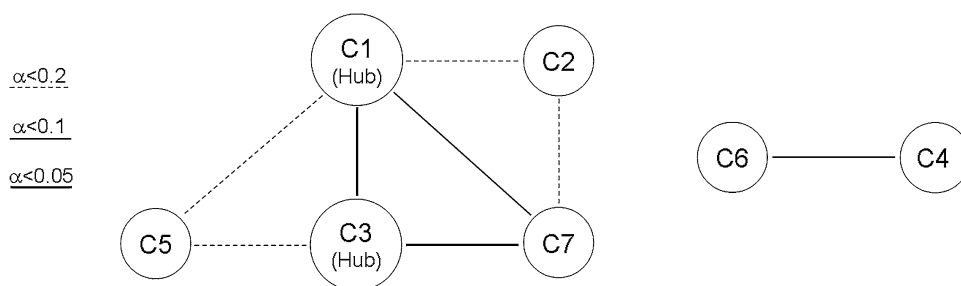
References

Beckmann CF, DeLuca M, Devlin JT, Smith SM (2005) Investigations into resting-state connectivity using independent component analysis. Phil Trans R Soc B 360:1001–1013

Bloomfield P (2000) Fourier analysis of time series. Wiley, New York

Broyd SJ, Demanuele C, Debener S, Helps SK, James CJ, Sonuga-Barke EJ (2008) Default-mode brain dysfunction in mental disorders: a systematic review. Neurosci Biobehav Rev 33(3): 279–296

Calhoun VD, Adali T, Pearlson GD, Pekar JJ (2001) A method for making group inferences from functional MRI data using independent component analysis. Hum Brain Mapp 14:140–151

Chen AC, Feng W, Zhao H, Yin Y, Wang P (2008) EEG default mode network in the human brain: spectral regional field powers. Neuroimage 41:561–574

Congedo M, Özen C, Sherlin L (2002) Notes on EEG resampling by natural cubic spline interpolation. J Neurotherapy 6(4):73–80

Congedo M, Gouy-Pailler C, Jutten C (2008) On the blind source separation of human electroencephalogram by approximate joint diagonalization of second order statistics. Clin neurophysiol 119:2677–2686

Damoiseaux JS, Rombouts SARB, Barkhof F, Scheltens P, Stam CJ, Smith SM (2006) Consistent resting-state networks across healthy subjects. Proc Natl Acad Sci USA 103(37):13848–13853

Fox MD, Raichle ME (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat Rev Neurosci 8(9):700–711

Greenblatt RE, Ossadtchi A, Pflieger ME (2005) Local linear estimators for the bioelectromagnetic inverse problem. IEEE Trans Sig Process 53(9):3403–3412

John ER, Ahn H, Prichep LS, Trepetin M, Brown D, Kaye H (1980) Developmental equations for the electroencephalogram. Science 210:1255–1258

Mantini D, Perrucci MG, Del Gratta C, Romani GL, Corbetta M (2007) Electrophysiological signatures of resting state networks in the human brain. Proc Natl Acad Sci USA 104(32):13170–13175

Nolte G, Bai O, Wheaton L, Mari Z, Vorback S et al (2004) Identifying true brain interaction from EEG data using the imaginery part of coherency. Clin Neurophisiol 115:2292–2307

Nunez PL, Srinivasan R (2006) Electric field of the brain, 2nd edn. Oxford University Press, New York

Pascual-Marqui RD (2007) Instantaneous and lagged measurements of linear and nonlinear dependence between groups of multivariate time series: frequency decomposition. arXiv:0711.1455v1

Pham D-T, Congedo M (2009) Least square joint diagonalization of matrices under an intrinsic scale constraint, ICA 2009. In: 8th international conference on independent component analysis and signal separation, March 15–18, Paraty, Brasil

Schmithorst VJ, Holland SK (2004) Comparison of three methods for generating group statistical inferences from independent component analysis of functional magnetic resonance imaging data. Magn Reson Imaging 19(3):365–368

van den Heuvel M, Mandl R, Hulshoff Pol H (2008) Normalized cut group clustering of resting-state FMRI data. PLoS ONE 3(4):e2001

Welch PD (1967) The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Trans Audio Electroacoustics 15(2):70–74