ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 22 (2004) 1023–1034

Concurrent EEG/fMRI analysis by multiway Partial Least Squares

Eduardo Martínez-Montes,[a,*] Pedro A. Valdés-Sosa,[a] Fumikazu Miwakeichi,[b] Robin I. Goldman,[c] and Mark S. Cohen[d]

[a] *Neurophysics Department, Cuban Neuroscience Center, Havana, Cuba*
[b] *Laboratory for Dynamics of Emergent Intelligence, RIKEN Brain Science Institute, Wako, Saitama, Japan*
[c] *Hatch Center for MR Research, Columbia University, New York, NY 10032, USA*
[d] *Ahmanson-Lovelace Brain Mapping Center, UCLA, Medical School, Los Angeles, CA 90095, USA*

Received 17 July 2003; revised 12 March 2004; accepted 17 March 2004

Data may now be recorded concurrently from EEG and functional MRI, using the Simultaneous Imaging for Tomographic Electrophysiology (SITE) method. As yet, there is no established means to integrate the analysis of the combined data set. Recognizing that the hemodynamically convolved time-varying EEG spectrum, S, is intrinsically multidimensional in space, frequency, and time motivated us to use multiway Partial Least-Squares (N-PLS) analysis to decompose EEG (independent variable) and fMRI (dependent variable) data uniquely as a sum of "atoms". Each EEG atom is the outer product of spatial, spectral, and temporal signatures and each fMRI atom the product of spatial and temporal signatures. The decomposition was constrained to maximize the covariance between corresponding temporal signatures of the EEG and fMRI. On all data sets, three components whose spectral peaks were in the theta, alpha, and gamma bands appeared; only the alpha atom had a significant temporal correlation with the fMRI signal. The spatial distribution of the alpha-band atom included parieto-occipital cortex, thalamus, and insula, and corresponded closely to that reported by Goldman et al. [NeuroReport 13(18) (2002) 2487] using a more conventional analysis. The source reconstruction from EEG spatial signature showed only the parieto-occipital sources. We interpret these results to indicate that some electrical sources may be intrinsically invisible to scalp EEG, yet may be revealed through conjoint analysis of EEG and fMRI data. These results may also expose brain regions that participate in the control of brain rhythms but may not themselves be generators. As of yet, no single neuroimaging method offers the optimal combination of spatial and temporal resolution; fusing fMRI and EEG meaningfully extends the spatio-temporal resolution and sensitivity of each method.

Keywords: N-PLS; EEG/fMRI fusion; PARAFAC; Multiway analysis; SSI; SITE

Introduction

The armamentarium of the neuroscientist now includes tools with spatial resolution ranging from centimeters to microns and temporal resolution from years to nanoseconds. Even so, no single tool provides an optimal combination of spatial and temporal resolution, and there generally exists a tradeoff in which improvement in one dimension of resolution requires compromises in the other (Churchland and Sejnowski, 1988). Extending our understanding of the functional architecture of the human brain necessarily requires a rational combination of multiple methods. Particularly attractive is the fusion of the superb temporal resolution of electroencephalography (EEG) or magnetoencepalography (MEG), with the excellent contrast and spatial resolving power of functional MRI (fMRI). Several methods of integration have been reported (Horwitz and Poeppel, 2002), each with its own approaches to analysis.

Under the assumption that the response of the brain to a set of stimuli or conditions is the same when acquired at different times, several groups (Babiloni et al., 2001; Baillet et al., 2001; Singh et al., 1998) have attempted the analysis of EEG and fMRI data, gathered separately. While this approach is not without problems (Gonzalez-Andino et al., 2001; Ioannides, 1999), there is increasing evidence that adequate modeling of multimodal data will allow the estimation of the underlying neural processes with simultaneously high spatial and temporal resolution (Trujillo et al., 2001).

More recently, methods have been described for the concurrent collection of EEG and fMRI data (Goldman et al., 2000). These methods make possible the study of dynamic relationship between fluctuations in the blood oxygenation level dependent (BOLD) signal and the properties of the electrical activity recorded on the scalp. Here, the fMRI and EEG data each necessarily provide evidence of the same underlying brain activity, although the extent to which they are measuring the same signals, or even signals from the same processes, is indeterminate.

In a method they have called Simultaneous Imaging for Tomographic Electrophysiology, or SITE, Goldman et al. (2002) created tomograms of the brain regions whose fMRI signal changes were associated with variations in alpha band power. In that work, 16 bipolar EEG channels were recorded under the eyes-closed resting state that is well known to produce elevated alpha wave activity. To match the EEG and the fMRI time courses, they then convolved the

* Corresponding author. Neurophysics Department, Cuban Neuroscience Center, Avenue 25, Esq. 158, #15202, PO Box 6412, 6414 Cubanacán, Playa, Havana, Cuba. Fax: +53-7-208-6707.
E-mail address: eduardo@cneuro.edu.cu (E. Martínez-Montes).
Available online on ScienceDirect (www.sciencedirect.com.)

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.03.038

measured alpha power at each time point with an a priori hemodynamic response model (Cohen, 1997) and calculated the correlation between the fluctuations of alpha activity and the BOLD time course at each voxel. Alpha activity was defined as the broad band spectral power in the frequency range 8–12 Hz calculated over the 2.5-s period needed to acquire each MRI volume and averaged over the occipital derivations (*T6-O2*, *O2-P4*, *T5-O1*, *O1-P3*). Positive correlations were found in thalamic voxels as well as in the insula, while negative correlations predominated in the parieto-occipital cortex. Thus, these correlation maps showed extended thalamocortical structures implicated in the generation of this EEG rhythm. A deeper analysis of these results, however, leads to further questions.

Traditionally, the EEG has been decomposed into a series of fixed broad spectral bands (delta, theta, alpha, beta, gamma, ...) based more on history and discovery than on a theoretical framework. This approach, although computationally convenient, may obscure the fact that the sources of each of these characteristic oscillations may or may not be unique (Szava et al., 1994). It has been shown that the EEG can be analyzed as a partially overlapping spectral components defined with high-frequency resolution (Pascual-Marqui et al., 1988); each spectral component being interpreted as reflecting the activity in a given oscillatory network. We seek here to associate each of these components with the BOLD-fMRI activity. In keeping with standard terminology in time–frequency decompositions (Chen et al., 2001), these components will be designated as "atoms".

While strong prior information suggests that the scalp locations best associated with alpha power fluctuations may well be near the occipital electrodes, other spectral components may have a more subtle or distributed relationship to scalp topography. Under these more general circumstances, it may be better to have a more data-driven means to estimate the linear combination of EEG measurements (or derivations) that correlate optimally with BOLD. Such estimates are likely to result in greater statistical power for the detection of EEG–fMRI relationships. Similarly, it might be desirable to look at the correlation of the EEG with a calculated optimum linear combination of all BOLD signals, rather than with each voxel separately.

Essentially, our goal has been to seek methods that best explain the spatio-temporal relationships between fMRI and the oscillatory components of the EEG without first forming a priori hypotheses as to which characteristics of the EEG are likely be of most interest. These considerations led us to search for methods of atomic decomposition of the EEG and methods for correlating the output of this decomposition with the fMRI data. There are a lot of well-known methods for data reduction of the EEG. Among them, Principal Components Analysis (PCA), Independent Components Analysis (ICA), and dictionary-based decompositions have been the most explored. They have been applied only to two-dimensional data. As the time-varying EEG spectrum is, in fact, a three-dimensional array (electrode pairs, frequencies, and time), it cannot be expressed conveniently as a matrix. The decomposition of such a multidimensional data has been better accomplished by a generalization of the Singular Value Decomposition known as Parallel Factor Analysis (PARAFAC) (Harshman, 1970), a tool that has been used previously in the analysis of evoked potentials (Field and Graupe, 1991) and pharmacological studies using high-dimensionality EEG data (Estienne et al., 2001). The most interesting advantage of the PARAFAC model is that it provides a unique decomposition without imposing orthogonality or independence constraints to the components. It is also valued for being a parsimonious and "easily interpretable" model (Bro, 1998).

Several calibration methods (Principal Components Regression, ridge regression) can be used for correlating the EEG decomposition and the fMRI data. Although some general guidelines have been given for establishing a hierarchy among them (Kiers, 1991), there is not definitive calibration method that one can stick to, since its correct application depends strongly on the behavior of the data considered. In a straightforward application of any of these methods (e.g., Principal Components Regression), one could use the EEG spectral power estimates (principal components of time-varying EEG spectrum), for different time segments, as the independent variable to be correlated with the fMRI. This procedure in two steps (decomposing and correlating) does not ensure that we are finding the optimal relationship between the EEG and the fMRI because decomposition is based on nonphysiological assumptions (e.g., it is unreasonable to expect that the activities of individual neural generators to be mutually orthogonal). Therefore, we should search for a method that is capable of simultaneously extracting EEG spectral components or atoms (and their scalp landscapes) having maximal temporal covariance with certain BOLD profiles. One possible candidate for such a multimodal analysis is Partial Least-Squares (PLS) regression, introduced in fMRI analysis by McIntosh et al. (1996). In PLS, the fMRI data are treated as a matrix (voxels by time). PLS identifies those linear combinations of fMRI voxels that have maximal temporal covariance with linear combinations of a second matrix of independent variables, measured at the same time points. The method hinges on calculating the Singular Value Decomposition (SVD) of the covariance matrix between the fMRI and independent variables. This method has been used for spatio-temporal analysis of event-related potentials (Lobaugh et al., 2001) and simultaneous EEG and MEG data (Düzel et al., 2003).

Fortunately, the PLS technique has been extended by Bro (1996) to deal with multidimensional data, obtaining a new model known as Multiway Partial Least Squares or just N-PLS. This model consists essentially of decomposing the independent and dependent data into multilinear models such that the score vectors from these models have pairwise maximal covariance. The multilinear decomposition is made in the same way as PARAFAC, thus inheriting both advantages and limitations of that model.

In this paper, the N-PLS model will be introduced for decomposing the EEG into a sum of atoms each with a specific spatial, temporal, and spectral factors or "signatures". Simultaneously, the fMRI data will be decomposed into the same number of atoms, each the product of spatial and temporal signatures, in such a way that the latter will have maximal covariance with the EEG temporal signature. The source localization of the EEG spatial signature (topography) of each atom will be examined, allowing separate analysis of the tomographic distribution of the EEG sources (what we will call sources of the "EEG rhythm") and those tomographic sources obtained as the fMRI tomograms that we interpret as the "brain rhythm" generating system. It should be noted that we have limited our consideration to oscillatory components of the EEG. While important, they do not exhaust the list of interesting phenomena that might possibly relate to the fMRI. Transient waveforms, for example, are not optimally described in the time–frequency framework. In principle, the methods developed here may be extended to consider this situation.

Methods

Consider a matrix, $\mathbf{F}_{(N_s \times N_t)}$, of the fMRI data ($N_s$ voxels, N_t time points) that is recorded simultaneously with the EEG time series from N_d electrodes. Further, define the EEG signal recorded during each TR (the period needed to collect an MRI volume) as a "segment." In the present case, the time-varying EEG spectrum, $\mathbf{S}(\omega)_{(N_d \times N_t)}$ (ω being the frequency), for N_t segments, was estimated via the Thomson multitaper method (Thomson, 1982). Let **s** be a reference EEG time signal, formed by selecting a linear combination, **a**, of the EEG electrode power in a given band of frequencies Ω, which was then filtered by the hemodynamic response, **H**:

$$\mathbf{s}_{(1\times N_t)} = \mathbf{a}^T_{(1\times N_d)} \sum_{\omega\in\Omega} \mathbf{S}(\omega)_{(N_d\times N_t)} \mathbf{H}_{(N_t\times N_t)} \quad (1)$$

where the symbol, $\mathbf{a}^T$, represents the transpose of vector **a**. Then, the correlations between the fMRI matrix and the reference EEG signal: $\mathbf{r}_{(N_s \times 1)}$= corr(**F**, **s**) are mapped.

In the analysis performed by Goldman et al. (2002), they chose an ad hoc linear combination, **a** (an occipital electrode set), and frequency band (8–12 Hz) for finding the reference EEG time signal. We will extend this analysis to estimate the optimal linear combination of electrodes, and a particular spectral window defining an optimal frequency band $\sum_{\omega\in\Omega} \mathbf{b}(\omega)\mathbf{S}(\omega)$. Finally, we will estimate a suitable linear combination, $\mathbf{u}^T_{(1 \times N_s)}$ of the elements of the fMRI matrix to be correlated with a particular EEG time signal.

Parallel Factor Analysis

Recognizing that the time-varying EEG spectrum may be expressed conveniently as a three-dimensional array makes possible the use of Parallel Factor Analysis (PARAFAC) (Carroll and Chang, 1970; Harshman, 1970), a generalization of Principal Component Analysis (PCA) for dealing with multidimensional data. With PARAFAC, the time-varying EEG spectrum is decomposed (in a least-squares sense) into trilinear components, or atoms, each being the product of a spatial, spectral and temporal factors, or signatures.

Unlike PCA, PARAFAC has no rotational freedom; therefore, the decomposition is unique, even without any orthogonality constraints. It has been shown that if the data are approximately trilinear, the correct number of components is used, and the signal-to-noise ratio is adequate, then the PARAFAC algorithm will show the true underlying phenomena (Kruskal, 1976, 1977). Moreover, PARAFAC provides a unique data-determined linear combination, i.e., a reference time signal, to correlate with the fMRI data. The use of PARAFAC in analyzing three-dimensional EEG data, (space, frequency, time) is described in a companion paper (Miwakeichi et al., 2004).

Then, applied to the time-varying EEG spectrum, which is expressed as a three-dimensional matrix $\mathbf{S}_{(N_d \times N_w \times N_t)}$, PARAFAC decomposition establishes an element-wise trilinear model for these data:

$$\hat{S}_{dwt} = \sum_{k=1}^{N_k} a_{dk} b_{wk} c_{tk} + \varepsilon_{dwt} \quad (2)$$

where d, w, and t designate electrode pairs, frequency, and time, respectively, and the term ε_{dwt} represents the error. The total number of components is N_k, each of which is designated by index k. Our problem is to find the so-called "loading matrices", **A**, **B**, and **C** whose N_k columns are the loading vectors $\mathbf{a}_{k(N_d \times 1)}$, $\mathbf{b}_{k(N_w \times 1)}$, and $\mathbf{c}_{k(N_t \times 1)}$ of elements a_{dk}, b_{wk}, and c_{tk} respectively.

We can fit the model expressed in Eq. (2) by finding

$$\min_{a_{dk} b_{wk} c_{tk}} \left|\left| S_{dwt} - \sum_{k=1}^{N_k} a_{dk} b_{wk} c_{tk} \right|\right|^2.$$

The interpretation of the loading vectors is as follows: $\mathbf{a}_k$ is the spatial signature of the kth atom, which is a representative topographic map, or linear combination of electrodes; $\mathbf{b}_k$ is the spectral signature for the kth atom and $\mathbf{c}_k$ is the temporal signature, or time course, for atom k. The only indeterminacies in the least-square solution are the order of components and the scaling of loading vectors. Thus, centering and scaling of the data are needed before decomposition, as is a convention for the signs and scale of the loadings. For PARAFAC, the resulting spectral and spatial loadings are normalized, while the non-normalized loading will be the temporal factor, reflecting the scale of the data.

It is important to select the most appropriate number, N_k, of components. The Core Consistency Diagnostic (Corcondia) is an approach for so doing that applies especially to PARAFAC models, and has been shown to be a powerful and simple tool for determining the appropriate number of components in multiway models (Bro, 1998). In this work, we use not only Corcondia but also the evaluation of the systematic variation left in the model's residuals.

PARAFAC has been extensively used in chemometrics, psychometrics, and econometrics. In the field of spatio-temporal analysis of Event-Related Potentials, PARAFAC has been shown to be formally equivalent to the Topographic Components Model (TCM) (Möcks, 1988a,b). Field and Graupe (1991) offered some general guidelines for the correct exploration of EEG data with PARAFAC. The basic pitfall of the application of PARAFAC is that the data are actually not trilinear, and, hence, a careful preprocessing and analysis of the results most be done for assessing the validity of the model.

Multiway Partial Least-Squares Regression

Despite being a useful tool for data explorations and to find a unique reference EEG time signal, the PARAFAC analysis leaves two important questions unanswered:

a) Which frequency components are related to the fMRI signal?
b) What is the optimal linear combination of EEG electrodes to correlate with the fMRI?

Partial Least-Squares regression is an automatic procedure to find the linear combination that maximizes the temporal correlation between the EEG and fMRI data (de Jong and Phatak, 1997; Martens and Naes, 1989). This method is similar to Principal Components Regression (PCR), where the independent variable is decomposed into a set of scores, and the dependent variable is regressed on these scores instead of the original variable. The main difference being that in PLS regression, both independent and dependent variables are decomposed such that these scores have maximal covariance; that is, the relevant variations of the independent variable for predicting the dependent variable are emphasized. An extension of the PLS regression model to three-way data was proposed by Ståhle (1989). Later, Bro (1996) developed a general multiway PLS (N-PLS) regression model that was shown

to be optimal according to the theory of PLS and had a particular case numerically equivalent to that of Ståhle. N-PLS seeks in accordance with the philosophy of PLS to describe the covariance of the dependent and independent variables. This is achieved by fitting multilinear models simultaneously for independent and dependent variables and for a regression model relating the two decomposition models. On the other hand, as covariance is the product of the correlation and the variances, these three measures actually are maximized collectively.

According to Bro (1996), the model is known as N-PLS or Multilinear PLS in general, and the specific model to be used in this work is called tri-PLS2. This follows from its having a three-way decomposition for the independent variable (*tri*), which will be the time-varying EEG spectrum, and a two-way or bilinear decomposition for the dependent variable (*2*), corresponding to the fMRI data. This can be considered as a form of PARAFAC decomposition constrained by additional conditions of maximal covariance with certain BOLD components. The structural model can be expressed as:

$$\hat{S}_{dwt} = \sum_{k=1}^{N_k} a_{dk} b_{wk} c_{tk} + e_{dwt}$$

$$\hat{F}_{st} = \sum_{k=1}^{N_k} u_{sk} v_{tk} + \varepsilon_{st}$$

where ε_{st} and e_{dwt} are elements of noise matrices and the index, s, represents the voxels or grid points inside the brain. These decomposition models are estimated iteratively, component-wise, by finding a set of normalized vectors, $\mathbf{a}_k$, $\mathbf{b}_k$, and $\mathbf{u}_k$ such that the least-squares score vectors, $\mathbf{c}_k$ and $\mathbf{v}_k$, have maximal covariance. It is worth underscoring that the N-PLS model is unique, as it consists of successively estimated one-atom models, each of which is itself always unique. On the other hand, note that the EEG data must first be preprocessed, both by removing muscle and motion artifacts, replacing them by linear interpolation of the data, and by convolution of the EEG spectrum with the hemodynamic impulse response function (Cohen, 1997). A graphical representation of the tri-PLS2 method is shown in Fig. 1.

The interpretation of loading vectors is straightforward. The spectral signature of the EEG for the k atom, $\mathbf{b}_k$, will allow the identification of those brain rhythms whose time-varying envelopes has maximal covariances with the BOLD signal. The spatial signature of the fMRI for atom k, $\mathbf{u}_k$, is a tomographic map (which is not a correlation map) showing those BOLD signals whose time courses are correlated maximally with the EEG. Finally, the spatial signature of the EEG, $\mathbf{a}_k$, is a representative topography of atom k, extracted by asking for the maximal temporal correlation between EEG and fMRI.

The decomposition is made component-wise; that is, for each component (atom), a rank-one model is built of both the independent variable 3D matrix **S**, and the dependent variable 2D fMRI matrix, **F**. These models are then subtracted from the original data, and a new atom of signatures is found from the residuals. The calculation for one atom of the tri-PLS2 model is developed in detail in Appendix A. As in PARAFAC, a convention about signs and scale is needed. In this case, the non-normalized factors will be the temporal signatures of both the EEG and the fMRI data, while the other signatures are normalized. Signs were assigned to ensure that the correlation between fMRI and EEG temporal signatures for the alpha atom were positive. Moreover, in this work, it is important to obtain smooth images as atoms of the spatial signature of the fMRI. For the sake of simplicity, the raw data can be presmoothed and the same smoothed signatures will be obtained from the decomposition (Bro, personal communication). Therefore, the raw fMRI data are presmoothed to obtain smoothed atoms for the spatial signature of fMRI. Smoothing consisted of applying the nearest neighbor moving average three times to the raw fMRI data.

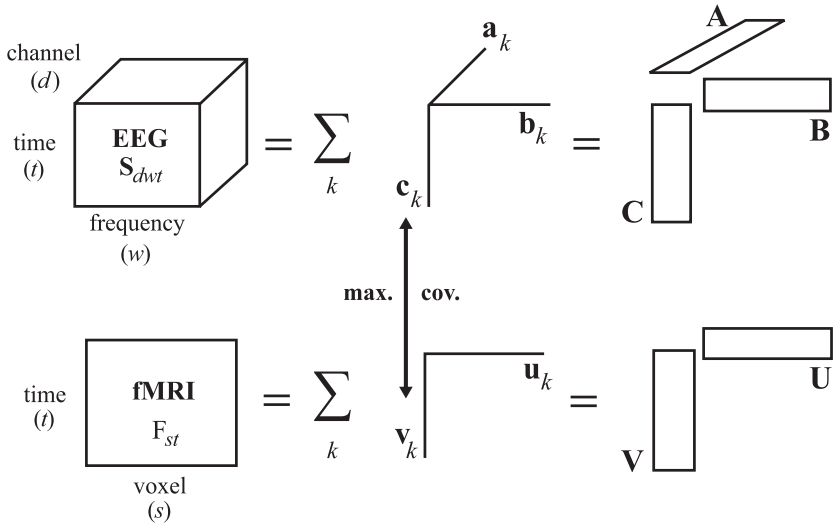

Fig. 1. Tri-PLS2 diagram. The time-varying EEG spectrum is represented as a three-dimensional array indexed by time (t), frequency (w), and channel (d). The fMRI matrix is indexed by time and voxels (s). Both data are decomposed into a sum of atoms or components. Each EEG atom have spatial ($\mathbf{a}_k$), spectral ($\mathbf{b}_k$), and temporal ($\mathbf{c}_k$) signatures. fMRI atoms have a spatial ($\mathbf{u}_k$) and temporal ($\mathbf{v}_k$) signature. The extraction of atoms is performed simultaneously by constraining the temporal signatures to have maximal covariance. Joining all atoms for each signature allows them to be expressed in matricial notation, obtaining corresponding matrices **A**, **B**, **C**, **V**, **U**.

Assessment of N-PLS model

The advantages of N-PLS over bilinear methods are that it is much more parsimonious, easier to interpret, and less prone to noise. This advantages hold even over nonlinear calibration models (e.g., feedforward neural networks) because they are bilinear in the decomposition of the independent variable and the nonlinearity is introduced only in the relation of this decomposition with the dependent variable. Another advantage is that the algorithm is faster than other multilinear decomposition methods (e.g., PARAFAC) due to the relatively few parameters to estimate and particularly, because the tri-PLS algorithm boils down to eigenvalue problems.

However, this model has its own pitfalls. The basic problem is the appropriateness of the trilinear model. As this is a data-dependent question, there is not a general and straightforward answer. If there is no any a priori knowledge about the three-way nature of a given data, one could try different methods to see which one describes the data best. In the case of several methods fitting the data equally well, one should choose the simplest model and in this regard multilinear models are preferred over bilinear ones. On the other hand, although it has been shown that models like N-PLS seldom fail to converge and offer degenerate solutions (Bro, 1998), these are problems that can arise in multiway methods and should be taken into account in the exploration of the data.

In practice, it is convenient to apply a PARAFAC decomposition to the EEG data before applying tri-PLS2 model. This initial exploration will allow to assessing the appropriateness of the trilinear model for the time-varying EEG spectrum, to identify possible outliers in the data, and the estimation of the number of significant atoms present in the data. The implementation of PARAFAC used in this work is contained in a Matlab Toolbox developed by Bro and available on the web. It provides several diagnostic tools, such as Corcondia, residuals plots, leverages plots, convergence, and explained variance of the data, among others.

As said above, the appropriate number of components was obtained with the residual analysis and the Corcondia index. This index was also used for assessing the trilinear structure of the data as shown in Estienne et al. (2001). The analysis of leverages allowed to detecting four outliers in the time mode. These four time windows (or segments) were discarded from the data for subsequent analysis. We also removed some constant signature (nonphysiologically meaningful) in the frequency mode by an adequate centering across this mode. Furthermore, comparison between the loadings of the time-varying EEG spectrum decomposition provided by PARAFAC and those provided by tri-PLS will validate (at a preliminary level) the truthfulness of the results obtained. A detailed explanation about the use of the diagnostic tools for this exploratory analysis and discussion of the reliability of PARAFAC model can be found in Bro (1998) and Miwakeichi et al. (2004).

Source localization analysis

The spatial signature for the time-varying EEG spectrum, $\mathbf{a}_k$, may be analyzed further by source reconstruction methods, such as Low-Resolution Electromagnetic Tomography (LORETA) (Pascual-Marqui et al., 1994) to find those underlying electrical sources that are correlated temporally with the BOLD signal. However, LORETA cannot be applied directly since $\mathbf{a}_k$ is not derived from voltages but rather from the power spectra of voltages. Therefore, in this case, we developed a procedure that allows the estimation of the spectra of the EEG sources on the basis of the spectra of the observed voltages. We shall call this type of source localization "Source Spectra Imaging" (SSI). This is based on the following assumptions:

- There is no spatial correlation between scalp voltage measurements.
- There is no spatial correlation between electric current densities inside the brain.
- The source spectra (variances of current densities in frequency domain) to be estimated will be the smoothest one in space.
- The source spectra is the same in the x, y, and z directions.

The detailed formulation for obtaining this inverse solution can be found in a companion paper (Miwakeichi et al., 2004). It must be emphasized that the assumptions behind this inverse solution can classify it as a distributed inverse solution, whose pitfalls and drawbacks have been extensively described in the literature (Fuchs et al., 1999; Pascual-Marqui, 1999).

Moreover, the EEG data analyzed in this work corresponds to voltages measured in an array of 16 bipolar pairs; therefore, to find the SSI solution, the problem of transforming these bipolar measurements into unipolar voltages must be addressed. We must thus construct the matrix $\mathbf{M}$ that transforms the spatial signatures of the EEG, $\mathbf{a}_k^{\text{uni}}$, obtained (ideally) from a unipolar array, into those measured from bipolar recordings (Eq. (3)). A partial representation of matrix $\mathbf{M}$ is given in Eq. (4). Then, $\mathbf{a}_k^{\text{uni}}$ is estimated by multiplying Eq. (3) by the Moore-Penrose pseudo-inverse of matrix $\mathbf{M}$.

$$\mathbf{a}_k = \mathbf{M}\mathbf{a}_k^{\text{uni}} \tag{3}$$

$$\boldsymbol{M} = \begin{array}{cccccccc} Fp2 & \dots F7 & F8 & \dots O2 & \dots T4 & \dots T6 & \\ 1 & \dots 0 & -1 & \dots 0 & \dots 0 & \dots 0 & Fp2-F8 \\ 0 & \dots 0 & 1 & \dots 0 \dots & -1 & \dots 0 & F8-T4 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots 0 & 0 \dots & -1 & \dots 0 \dots & 1 & T6-O2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \vdots \end{array} \tag{4}$$

Further, with this method, we can visualize the spatial signatures of the EEG obtained by tri-PLS2 decomposition (which would correspond to the bipolar topographies) as a topographic map on the head. Finally, it is noteworthy that these topographic maps, and their SSI solutions, are essentially dimensionless, as the former is normalized as part of the scale convention for the tri-PLS2 model.

Statistical inference

Our first inferential problem is to determine whether there is a significant correlation between the time courses of the EEG and fMRI. This can be tested readily by permutation of the time

segments in the time-varying EEG spectrum, which will destroy any temporal correlation between the EEG and fMRI data (Galán et al., 1997). This procedure is not appropriate, however, if there is any autocorrelation in the time series of the EEG data. Using the ARFIT toolbox for Matlab (Schneider and Neumaier, 2001), we fitted an autoregressive model of order 2 (selected automatically by Schwarz's criterion) to the time course of time-varying EEG spectrum. With this information, we applied a block bootstrap method, which is adequate in the case of weak dependence of observations (time points in this case). The method consists of resampling with replacement, using blocks of consecutive time points instead of individual time points. The length of the blocks was chosen to be great enough to preserve the original dependence, so that the empirical distribution of statistics for blocks will resemble that for the original time points (Davison and Hinkley, 1997). On the other hand, it is also desirable to have as many blocks as possible. In our case, we use nonoverlapping blocks of length $l = 2p+1$; $p = 2$ being the order of the autoregressive model. Thus, by applying the tri-PLS2 method for N resampled series, we obtained N different decompositions into atoms of corresponding signatures for each modality. The correlation coefficients between the EEG and corresponding fMRI time courses for each atom were then computed. From the 95th percentile of the empirical distribution of these correlations, we established a significance level for testing of the original correlation.

Our second inferential problem is to determine which voxels in the spatial signature of the fMRI are significantly different from zero. This is important for identifying brain regions that contribute to a particular EEG-fMRI temporal correlation. Thus, for this problem, we used a simple jackknife resampling procedure (Davison and Hinkley, 1997) from which a pseudo t image was constructed. In this specific case, the jackknifed estimate was obtained as follows: the leave-one-out spatial signatures ($\mathbf{u}_i$; $i = 1 \ldots N_t$) of the fMRI were created by leaving out time points one at a time and applying the tri-PLS2 model to the truncated data. The jackknife pseudo observations were then computed as:

$$\boldsymbol{u}_i^* N_t \boldsymbol{u} - (N_t - 1)\boldsymbol{u}_i; \quad i = 1 \ldots N_t$$

where $\mathbf{u}$ is the fMRI spatial signature corresponding to the complete data. This equation holds for all components although we have eliminated the subscript k for simplicity. Using the mean of the pseudo observations ($\bar{\boldsymbol{u}}^* = \frac{1}{N_t}\sum_{i=1}^{N_t} \boldsymbol{u}_i^*$) and the standard deviations $\left(\sigma_{\boldsymbol{u}^*} = \sqrt{\frac{1}{N_t}\sum_{i=1}^{N_t}(\boldsymbol{u}_i^* - \bar{\boldsymbol{u}}^*)^2}\right)$, the pseudo $\mathbf{t}$ image for each atom can be computed as $\mathbf{t}_{\text{image}} = \sqrt{N_t}\frac{\bar{\boldsymbol{u}}^*}{\sigma_{u^*}}$.

Experimental data

The EEG was sampled at 200 Hz from an array of 16 bipolar pairs, (*Fp2-F8*, *F8-T4*, *T4-T6*, *T6-O2*, *O2-P4*, *P4-C4*, *C4-F4*, *F4-Fp2*; *Fp1-F7*, *F7-T3*, *T3-T5*, *T5-O1*, *O1-P3*, *P3-C3*, *C3-F7*, *F7-Fp1*), with an additional channel for the EKG and scan trigger. The fMRI time series was measured in six slice planes (4 mm, skip 1 mm) parallel to the AC–PC line, with the second from the bottom slice through AC–PC. More details about this data set can be found in Goldman et al. (2002). In the work presented here, we have analyzed five simultaneous EEG/fMRI recordings from three different subjects. Informed consent was obtained from all volunteers based on a protocol approved previously by the UCLA Office for the Protection of Research Subjects.

Results

Both PARAFAC and N-PLS techniques were applied to the recorded data sets, and yielded similar results for all subjects. There was no statistical inference about differences among subjects, so, for the purpose of this paper, we present representative data from a single subject. As a first exploration of the data, a PARAFAC model was fitted to the time-varying EEG spectrum. The appropriate number of components for this model was chosen using Corcondia (see above). The model was fitted using direct trilinear decomposition for its initial values.

Three significant atoms or components, characterized by their spectral signature, were extracted by PARAFAC (Fig. 2A). It is

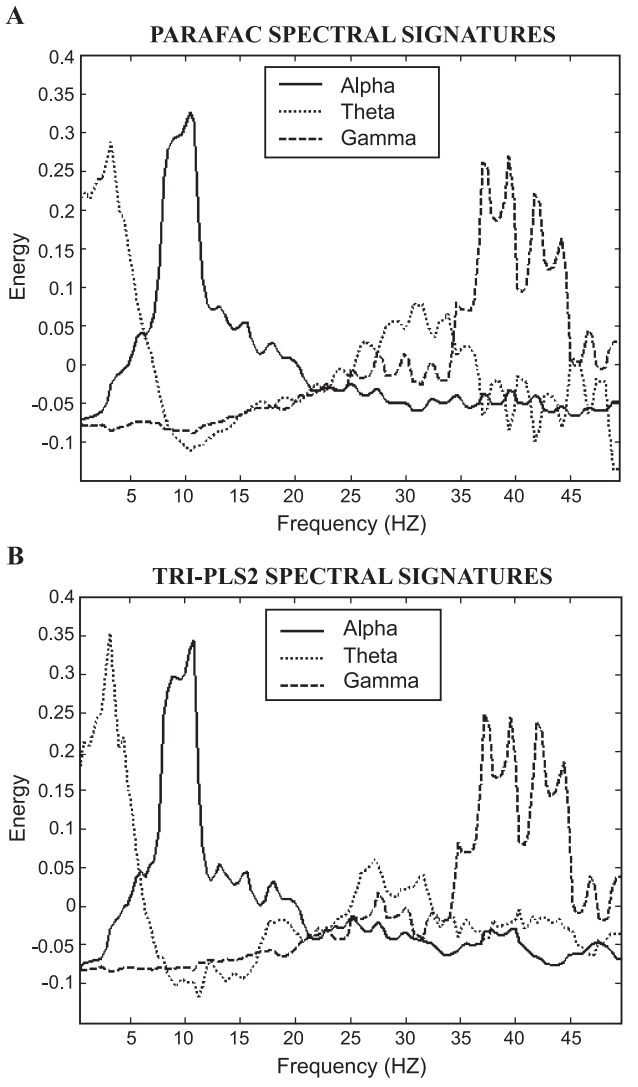

Fig. 2. Spectral signatures of the EEG decomposition. (A) Spectral signatures obtained from PARAFAC decomposition of the time-varying EEG spectrum. Three atoms were extracted. The first has a spectral peak around 10 Hz, corresponding to the well-known alpha rhythm. The second has a spectral peak around 4 Hz, which is a value usually assigned to theta activity. The third atom corresponds to a fast activity with spectral peaks from 35 to 45 Hz, in the gamma range. (B) Spectral signatures of the time-varying EEG spectrum, obtained from the tri-PLS2 model. Three atoms or components were extracted. These spectra resemble strongly those obtained from PARAFAC decomposition.

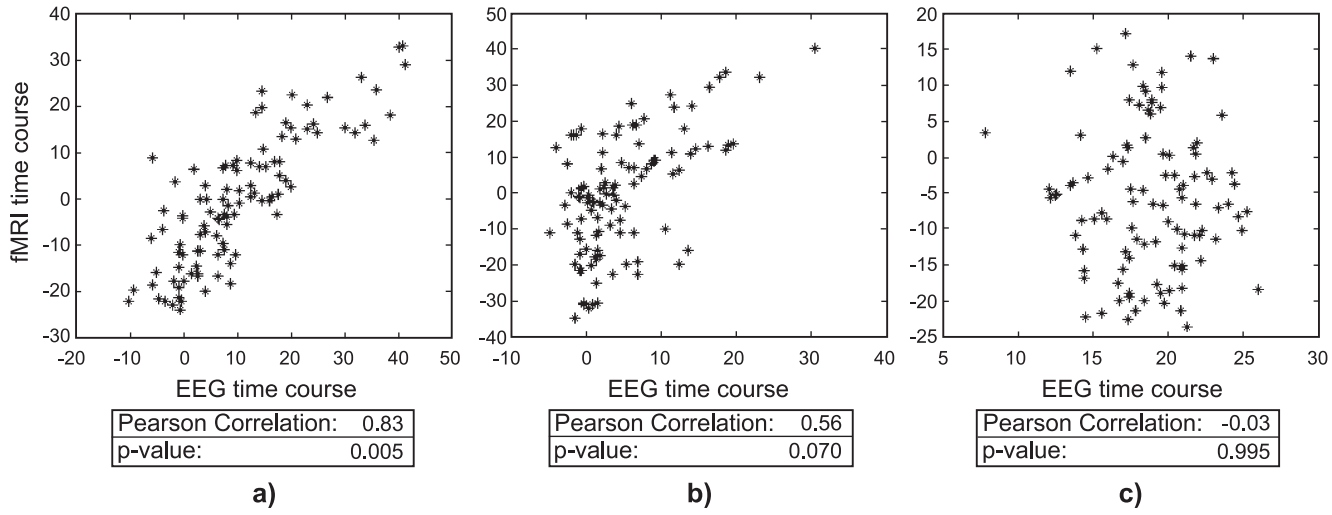

Fig. 3. Scatter plot of fMRI temporal signature against EEG temporal signature. (a) Alpha atom. A nearly linear positive dependence can be seen. The Pearson correlation value is 0.83, corresponding to $P = 0.005$. (b) Theta atom. The linear dependence between the EEG time course and fMRI time course has a positive correlation value of 0.56. However, it is not significant, $p = 0.07$. (c) Gamma atom. There is no clear linear dependence.

easy to recognize the alpha atom with its peak near 10 Hz. A slower theta activity peak is also present with a maximum around 4 Hz, as is a gamma peak in the range from 35 to 45 Hz. The Corcondia for this fit was around 93%, and the explained variation of the data was 53.5%. Moreover, PARAFAC allowed identification of outliers in the temporal signature, which were eliminated from the data for subsequent PARAFAC and posterior analyses.

Using this information, the N-PLS model was applied for only three atoms. In Fig. 2B, the spectral signatures for all atoms are shown, and they resemble strongly the spectra found by PARAFAC decomposition. Fig. 3 shows scatter plots of the temporal signatures of the fMRI vs. the EEG separately for each atom. The alpha and theta activities seem to have clearly positive correlations, but gamma activity does not. The Pearson correlation values are shown for each activity band. Supporting the visual impression, correlations were highest for the alpha atom. By using the 1000 samples of the block bootstrap test described previously, only the alpha atom presents a correlation value with probability lower than 0.05. The theta atom has a non-negligible correlation value, whose empirical probability is slightly higher than the predetermined theoretical significance level of 0.05.

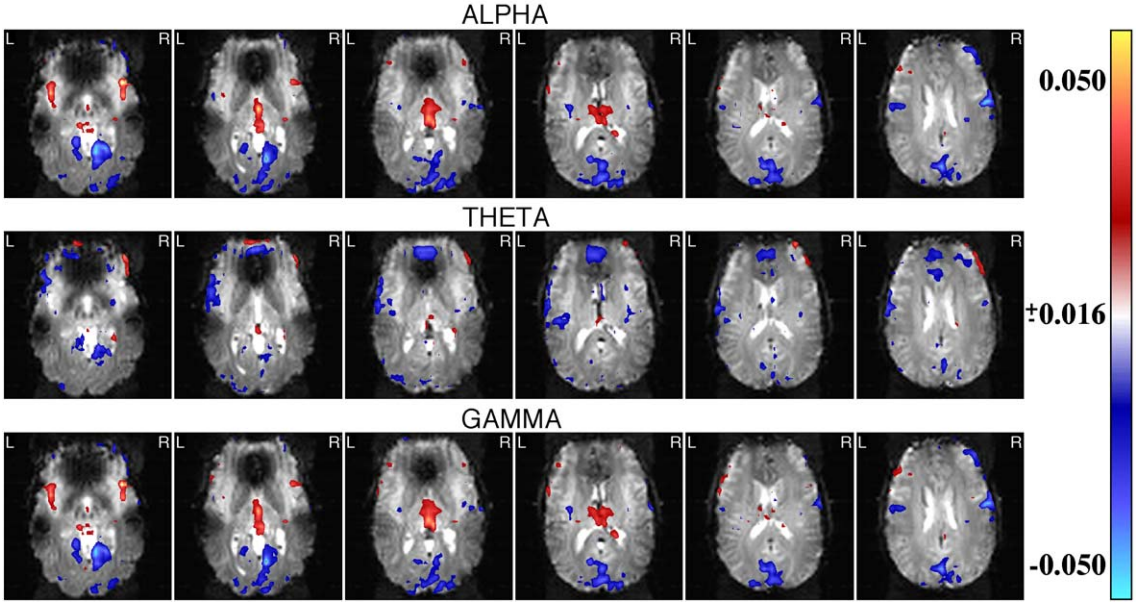

Fig. 4. fMRI spatial signatures for the three atoms. All images were plotted following a color scale from −0.05 to 0.05. However, the components have different minimum and maximum values. Alpha and Gamma atoms have a maximum of 0.055 and a minimum value of −0.066. Theta atom has a maximum value of 0.037 and a minimum of −0.066. The threshold was chosen conveniently to 0.016 for better visualization of the areas with higher values.

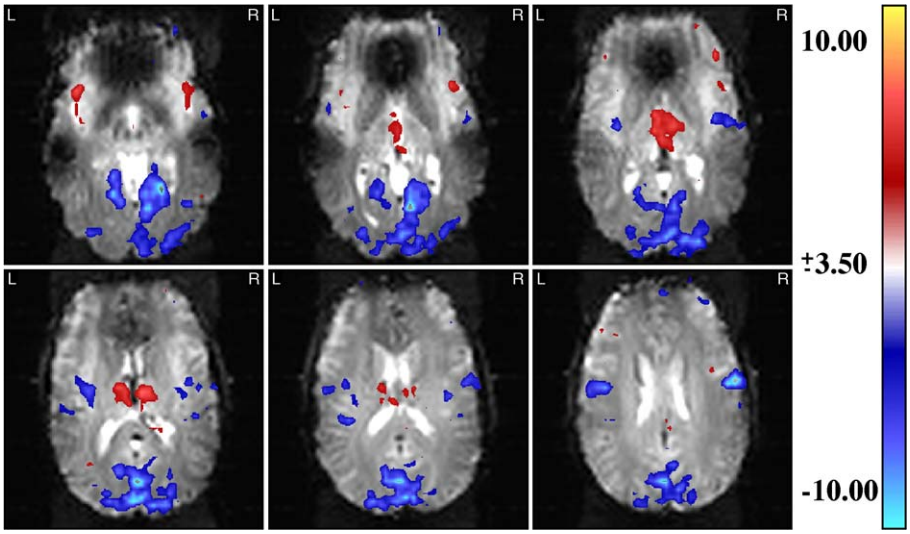

Fig. 5. Jackknifed pseudo t image for the fMRI spatial signature of the alpha rhythm atom. The jackknife procedure consisted of leaving out temporal points one at a time and applying the tri-PLS2 model to the truncated data. Then, a t value was calculated for each voxel and the resulting image was thresholded to a significance value of ± 3.5. Blue regions (anterior median occipital, lateral occipital, occipital pole, and left and right temporal superior) represent those areas with significantly negative temporal correlation with EEG. Thalamus and insula are red representing a significant positive correlation between EEG and fMRI time courses.

Fig. 4 shows the spatial signature of the fMRI decomposition: the $\mathbf{u}_k$ vectors. These are shown as tomograms in which those regions that have negative temporal correlation between EEG and fMRI are blue and those that have positive temporal correlation appear in red. For the alpha atom, the fMRI spatial signature shows positive activation of thalamus and insula, while occipital and superior temporal regions are activated negatively. The theta atom showed predominantly negative activation of anterior cingulate and occipital regions, while the gamma atom resembles the alpha component. For testing the robustness of this type of image, a pseudo t image of the alpha atom was calculated, it being the only atom having a significant temporal correlation with the EEG. This image is shown in Fig. 5, and was achieved by the jackknife procedure described above. In this figure, blue regions, (anterior median occipital, lateral occipital, occipital pole, and left and right temporal superior) represent those areas with significant negative

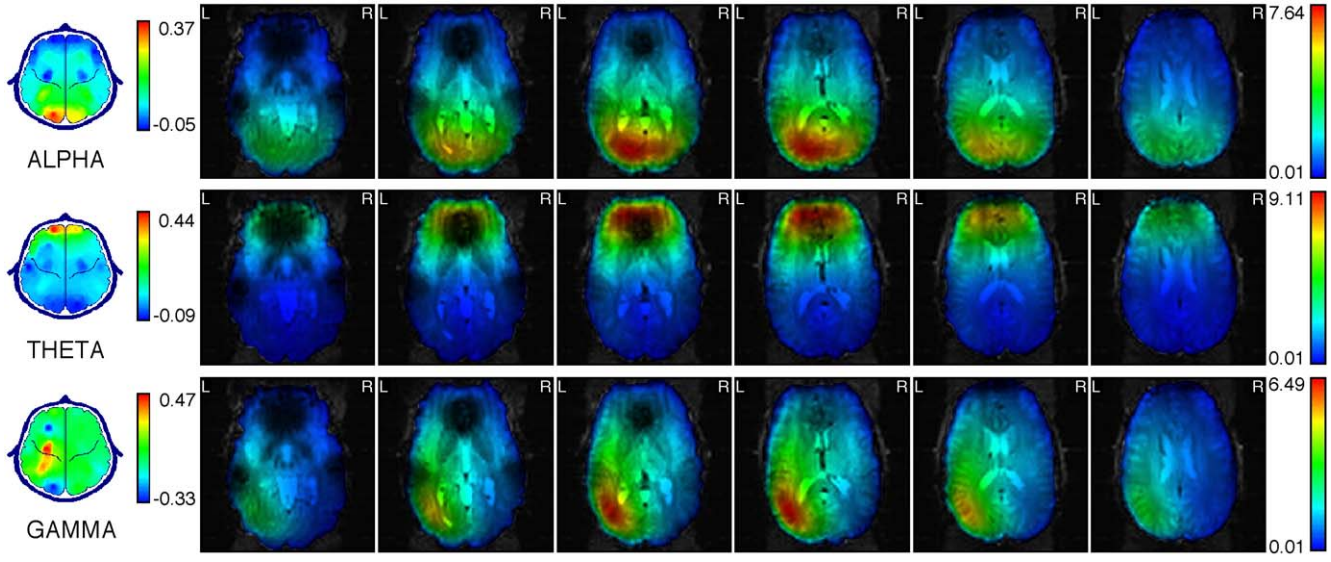

Fig. 6. Spatial signatures of the EEG and its SSI solutions. The topographical representation of spatial signatures of the EEG is shown at the far left. This map was calculated by pseudo-inverting the matrix that transforms topographies from unipolar recordings into those obtained with bipolar derivations. The alpha atom shows higher values at posterior regions, the theta topography has higher values located in frontal regions, and the gamma atom shows maximum values in the left parieto-temporal area. The corresponding SSI solutions are to the right. Maximum activation for the alpha component is located in the occipital area, with higher activation in the left hemisphere. Theta sources are in the anterior cingulate region, and the activated region for the gamma atom is located in the parieto-temporal area. Units for inverse solutions are ignored because energy values for the topographies are plotted and have been normalized as part of the N-PLS algorithm.

temporal correlation with the EEG. The areas corresponding to thalamus and insula are red, representing a significant positive correlation between EEG and fMRI time courses.

From the spatial signature, $\mathbf{a}_k$, of the time-varying EEG spectrum, we estimated those regions inside the brain that contribute to the EEG and that are correlated temporally with fMRI. Fig. 6 shows the topographies or EEG spatial signatures, and their corresponding SSI solutions (current density spectra) for each atom. The topography of the alpha atom shows higher values at posterior regions; the theta topography has higher values in frontal regions, and the gamma atom shows maximum topographic values in the left parieto-temporal area. The Source Spectra Imaging solution for the alpha component showed its maximum activation in the occipital area, with higher activation in the left hemisphere. Sources for theta atom are in the anterior cingulate region, and the activated region for the gamma atom is in the parieto-temporal area.

Discussion

This paper introduces a new method, trilinear Partial Least Squares (tri-PLS2), for the analysis of concurrent EEG/fMRI recordings. This is the first use of Partial Least-Squares techniques to carry out multimodal neuroimaging fusion. Our objective is to identify the coherent systems of neural oscillators that contribute to the spontaneous EEG. Doing so requires the solution of three related problems: (i) decomposing the EEG, in the space–frequency–time domain, into a set of components or atoms, (ii) establishing the relation of these EEG components to concurrent BOLD fluctuations, and (iii) analyzing the sources of the EEG atoms. We shall consider each of these problems in turn. At the outset it should be stressed that the two phenomena—EEG and BOLD—evolve over very different time scales. In fact, we shall be analyzing the *evolutionary spectrum* (Priestley, 1965) of the EEG, a concept based on a locally stationary modeling of the electroencephalogram (Dahlhaus, 1997). It is only the *envelope* of the waves usually analyzed by electroencephalography that will be matched to BOLD.

Atomic decomposition of the EEG

The analysis of the *evolutionary spectrum* of the EEG produces a three-dimensional data array (space–frequency–time). The first choices that come to mind for the decomposition of this array are either Principal Components Analysis (PCA) or Independent Components Analysis (ICA), a set of techniques that have received much recent attention. We decided to avoid these methods, however, for two reasons: First, they achieve a unique decomposition into atoms only by imposing arbitrary mathematical constraints (orthogonality and independence, respectively), and second, these methods are targeted toward two-dimensional arrays (matrices). In our situation, this means "unfolding" the data, stacking the time and frequency components along one dimension, and thereby destroying their distinction; keeping these different dimensions separate seems a much better alternative. A first attempt at a space–frequency–time atomic decomposition was reported in a paper by Koenig et al. (2001). In their method, the decomposition is carried out in several stages; first by the identification of time–frequency atoms, and then by the estimation of distinct topographies that are stable over time. This separation into two stages of analysis is not conceptually necessary, and in fact is not optimal.

Our trilinear method based on Parallel Factor Analysis, introduced in the present paper, allows space–frequency–time estimation in a single step, by minimization of an explicit objective function. The resulting decomposition is intrinsically unique and specifies atoms that are defined as spectral components that vary over time and have a specific topography. A more detailed description of the combined use of PARAFAC, and distributed inverse solutions for in vivo imaging of neural oscillatory systems, is the subject of a companion paper (Miwakeichi et al., 2004). A consistent finding in all data sets analyzed was the appearance of three components whose peaks were within the traditional theta, alpha, and gamma bands. Thus, when looking at the relations of the EEG with BOLD, it is potentially important not to constrain the analysis to a single frequency band, as was done by Goldman et al. (2002), although the present data do not show strong fMRI correlation with the EEG signal in the other bands.

It is remarkable that the restriction of maximal correlation with the BOLD signal produces spectra that are practically the same as those obtained by the PARAFAC decomposition. Based on the diagnostic tools, the physiological interpretability, and the replicability among several data sets, we can say that these are meaningful results, although we cannot ensure that they correspond with the real underlying physical phenomena. Therefore, it can be concluded that we have obtained robust and physiologically meaningful results with the use of tri-PLS algorithm.

Relating EEG atoms to the BOLD signal

As shown here, it is possible to constrain the trilinear EEG atomic decomposition further by requiring maximal temporal correlation with the BOLD signal, a procedure that extends the classical Partial Least-Squares technique. It is important to say that the correlation found between both temporal signatures was assessed by a block bootstrap method, which is a strong diagnostic tool for obtaining reliable results. Furthermore, the spatial signature of the fMRI was also statistically validated with the use of a jackknife procedure. These kind of diagnostic tools provide additional evidence on the robustness of the model assumed, i.e., about how well the properties of the data fit the assumptions of the model.

The alpha component has a temporal relation to the BOLD signal that is significant. The regional distribution of the fMRI spatial factor corresponds closely to that described by Goldman et al. (2002), for alpha activity, thus confirming their conclusions. Since the correlation between the EEG and BOLD temporal factors are positive, it becomes clear that the image shown in Fig. 5 is equivalent to the correlation map presented by that group. In particular, there is a positive relation between thalamic and insular BOLD activity and the EEG time course for alpha component. On the other hand, the BOLD signals within parieto-occipital and somatosensory cortices are related inversely to EEG. This latter negative correlation is probably due to a decrease in the amplitude of the EEG in activated cortex in this band, resulting from the temporal resynchronization of the postsynaptic potentials of the involved neural circuits.

The extracted theta component showed moderate temporal correlations that did not reach the pre-established 0.05 level of statistical significance. An examination of the spatial distribution

of the fMRI spatial signature for this atom shows a frontal activation. It is tempting to speculate that this component corresponds to a frontal midline theta rhythm that has not been adequately resolved due to the limited spatial coverage of the brain by the fMRI protocol used. The gamma component was not correlated with the recorded BOLD signal. Once again, we cannot exclude the possibility that better spatial coverage of the brain might reveal such correlations. Further, we can speculate that gamma fluctuations might relate to dynamic and transient assemblies of systems of brain activation (Tallon-Baudry and Bertrand, 1999) that are not stable throughout the recording period.

Analyzing the sources of the EEG atoms

A strength of both PARAFAC and tri-PLS2 is that they identify definite topographic patterns that can be subjected to source localization. These inverse solutions interpreted together with the fMRI spatial factors provide new information on the sources of EEG rhythms.

The Source Spectra Imaging solution for the alpha component reveals activation predominantly in the parieto-occipital region. This corresponds with results on the origins of alpha rhythm that have been reported previously, both using a frequency domain dipole solution (Valdés-Sosa et al., 1998) as well as frequency domain distributed solutions (Casanova et al., 2000). An interesting fact is that the thalamus shows very little activation, in contrast to the high positive correlation found by Goldman et al. (2002), and confirmed by the tri-PLS2 fMRI spatial signature.

This dissociation between the sources of the spatial signature of the EEG atoms and the spatial signature of the fMRI of the alpha atom is likely due to the negligible contribution of the primary current sources of thalamic neurons to the scalp EEG. In this case, the observed correlations between thalamic BOLD and EEG must be indirect. For example, the thalamus is probably correlated negatively with the parieto-occipital cortex, which seems to be the location of the generators of the ''EEG alpha rhythm''. Because the BOLD signal in this region is also correlated negatively with the alpha EEG spectrum, this would explain a positive correlation between alpha power and thalamic metabolic activity as an indirect effect through parieto-occipital cortex. In the terminology of Friston et al. (1996), there is a functional connectivity between the EEG and thalamus, but the effective connectivity path would not be direct, being mediated instead by the parieto-occipital cortex. Thus, according to the definitions given here, insula, thalamus, and parieto-occipital cortex are generators of the ''alpha brain rhythm'' while only parieto-occipital cortex contribute to the ''EEG alpha rhythm''. We note that the analyses presented in this paper do not allow the distinction of whether a structure belonging to a rhythm generating system oscillates in that frequency range. It seems unlikely that joint EEG/fMRI recordings can resolve the extent of phasic, versus tonic, participation in a brain rhythm of a structure that does not produce a measurable EEG. In other words, there is still invisible information for an EEG/fMRI fusion analysis, namely, the fine temporal characteristics of those areas that are invisible in the scalp EEG. In future planned experiments, it may be possible to resolve this issue through conjoint fMRI and depth electrode studies.

The tri-PLS2 method introduced in this paper is an example of multimodal image fusion, which takes advantage of the spatial resolution of the fMRI, as well as the temporal resolution of the EEG. This data analytic approach is capable of parsimoniously determining which EEG components are significant in the final analyses, and of revealing new features of the data by differentiating regions exposed within the fMRI data from those indicated solely through inverse solutions using the EEG. We are pursuing a number of improvements to enhance the integration of both types of data modalities by this method. In the first, we are developing a variant of tri-PLS2 that will estimate the spatial components, not on the scalp topography, as is done now, but instead directly in the source space. This would integrate source localization into the procedure rather than applying it as a postprocessing step for the topographies of the EEG atoms. Additionally, the autocorrelation of both the EEG and BOLD time series will be taken into account, whereas the model presented here ignores this information. Finally, it may well be that there are interactions between time, topography, and frequency spectrum that the current algorithm cannot account for.

Acknowledgments

The authors want to thank Dr. Eduardo Aubert from the Cuban Neuroscience Center for his important collaboration on handling MRI images used in this paper. Also, for their support and continuous comments on this work, we thank Nelson Trujillo, Lester Melie, and Ernesto Palmero from the Neurophysics Department of the Cuban Neuroscience Center. MSC is supported for this work under NIH DA15549 and DA13054.

Finally, we appreciate and thank Prof. Yoko Yamaguchi, head of the Laboratory for Dynamics of Emergent Intelligence of the RIKEN Brain Science Institute, for her hospitality and support for the successful conclusion of this work.

Appendix A. Tri-PLS2 algorithm

To calculate an atom of the tri-PLS2 model, we rewrite the model for dependent and independent variables taking only one atom, k, into account. Here the independent variable is the time-varying EEG spectrum, convolved previously with the hemodynamic response function, which is a three-way array $\mathbf{S}$. The dependent variable is the fMRI 2D matrix $\mathbf{F}$. The structural models then are

$$\hat{S}_{dwt} = a_{dk} b_{wk} c_{tk} \tag{A.1}$$

and

$$\hat{F}_{st} = u_{sk} v_{tk}. \tag{A.2}$$

The score vectors are those dependent on time (temporal signatures), i.e., $\mathbf{c}_k = (c_{1k},\ldots,c_{tk},\ldots,c_{N_{tk}})^T$ and $\mathbf{v}_k = (v_{1k},\ldots,v_{tk},\ldots,v_{N_{tk}})^T$; the others are also called the weights (spatial and spectral signatures of the EEG, spatial signature of the fMRI). The indices $t = 1,\ldots, N_t$, $w = 1,\ldots, N_w$, $d = 1,\ldots, N_d$ and $s = 1,\ldots, N_s$ represent time, frequency, channels, and voxels, respectively. For given weight vectors, the least-squares solution for determining the score vectors are:

$$c_{tk} = \sum_{w=1}^{N_w} \sum_{d=1}^{N_d} S_{dwt} a_{dk} b_{wk} \tag{A.3}$$

and

$$v_{tk} = \sum_{s=1}^{Ns} F_{st} u_{sk}. \tag{A.4}$$

Our problem is to find a set of normalized weight vectors, $\mathbf{a}_k$, $\mathbf{b}_k$, and $\mathbf{u}_k$, which produce score vectors, $\mathbf{c}_k$ and $\mathbf{v}_k$, having maximal covariance. The objective function to be maximized is:

$$\max_{\mathbf{a}_k,\mathbf{b}_k,\mathbf{u}_k} \left[\sum_{t=1}^{N_t} c_{tk} v_{tk} \middle| c_{tk} = \sum_{w=1}^{N_w} \sum_{d=1}^{N_d} S_{dwt} a_{dk} b_{wk} \wedge v_{tk} = \sum_{s=1}^{N_s} F_{st} u_{sk} \right] \tag{A.5}$$

For simplicity, the restriction of normalization on the weight vectors is not made explicit. Eq. (A.5) is not strictly correct because there is no correction for degrees of freedom, but as this correction is constant for a given atom, it will not affect the maximization. Eq. (A.5) also does not express the covariance if **S** and **F** have not been centered.

The next procedure is performed in two ways. First, Eq. (A.5) could be taken to:

$$\max_{\mathbf{a}_k,\mathbf{b}_k} \left[\sum_{t=1}^{N_t} \sum_{w=1}^{N_w} \sum_{d=1}^{N_d} S_{dwt} v_{tk} a_{dk} b_{wk} \right] = \max_{a_k,b_k} \left[\sum_{w=1}^{N_w} \sum_{d=1}^{N_d} z_{dwt} a_{dk} b_{wk} \right], \tag{A.6}$$

where $z_{dwk} = \sum_{t=1}^{N_t} S_{dwt} v_{tk}$ are the elements of an auxiliary matrix, $\mathbf{Z}_k$. If one writes Eq. (A.6) in matrix notation, the equation will become:

$$\max_{\mathbf{a}_k,\mathbf{b}_k} [\mathbf{b}_k^T \mathbf{Z}_k \mathbf{a}_k] \Rightarrow (\mathbf{b}_k, \lambda_k, \mathbf{a}_k) = \mathrm{SVD}(\mathbf{Z}_k, 1). \tag{A.7}$$

In other words, the weight vectors, $\mathbf{a}_k$ and $\mathbf{b}_k$ can be computed from the first component of a singular value decomposition of $\mathbf{Z}_k$ [SVD($\mathbf{Z}_k$, 1)]. This follows directly from the properties of SVD.

Second, substituting in Eq. (A.5) the corresponding score vector for the dependent variable:

$$\max_{\mathbf{u}_k} \left[\sum_{t=1}^{N_t} \sum_{s=1}^{N_s} F_{st} c_{tk} u_{sk} \right] = \max_{u_k} \left[\sum_{s=1}^{N_s} y_{sk} u_{sk} \right], \tag{A.8}$$

where $y_{sk} = \sum_{t=1}^{N_t} F_{st} c_{tk}$ are the elements of an auxiliary vector $\mathbf{y}_k$. Since $\mathbf{u}_k$ is restricted to be normalized, the maximum value of the expression (A.8) is reached when $\mathbf{u}_k$ is a unit vector in the same direction as $\mathbf{y}_k$. Therefore, the solution is:

$$\mathbf{u}_k = \frac{\mathbf{y}_k}{\|\mathbf{y}_k\|} = \frac{\mathbf{F}^T \mathbf{c}_k}{\|\mathbf{F}^T \mathbf{c}_k\|} \tag{A.9}$$

On the other hand, through the models of the data sets given in Eqs. (A.1) and (A.2), the prediction model between **S** and **F** is found by using a regression model for the so-called inner relation (established for the loadings matrices, i.e., for all atoms at the same time):

$$\mathbf{V} = \mathbf{CX} + \mathbf{E}_v.$$

This expression ensures that the maximum covariance restriction holds, and allows prediction of new samples of dependent variables. As the different atoms for score vectors are not always orthogonal, all of these atoms must be taken into account in calculating regression coefficients. The regression thus leads to:

$$\mathbf{x}_k = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{v}_k. \tag{A.10}$$

Finally, we can summarize the algorithm as follows:

1. Center **S** and **F**.
2. Let $\mathbf{v}_k$ equal a column in **F**.
3. Atom $k = 1$.
4. Compute matrix $\mathbf{Z}_k$ using **S** and $\mathbf{v}_k$.
5. Determine $\mathbf{a}_k$ and $\mathbf{b}_k$ from Eq. (A.7).
6. Calculate $\mathbf{c}_k$ from Eq. (A.3).
7. Compute $\mathbf{u}_k$ from Eq. (A.9).
8. Compute $\mathbf{v}_k$ from Eq. (A.4).
9. If the results converge, continue. Otherwise go to step 4.
10. Do the regression, finding $\mathbf{x}_k$ from Eq. (A.10).
11. $\mathbf{S}_t = \mathbf{S}_t - c_{tk} \mathbf{b}_k \mathbf{a}_k^T$ (for all t) and $\mathbf{F} = \mathbf{F} - \mathbf{C}\mathbf{x}_k \mathbf{u}_k^T$.
12. $k = k + 1$. Repeat from 4 until **F** is properly described.

References

Babiloni, F., Babiloni, C., Carducci, F., Angelone, L., Del Gratta, C., Romani, G.L., Rossini, P.M., Cincotti, F., 2001. Linear inverse estimation of cortical sources by using high resolution EEG and fMRI priors. IJBEM 3, 1.

Baillet, S., Leahy, R.M., Singh, M., Shattuck, D.W., Mosher, J.C., 2001. Supplementary motor area activation preceding voluntary finger movements as evidenced by magnetoencephalography and fMRI. IJBEM 3, 1.

Bro, R., 1996. Multi-way calibration. Multi-linear PLS. J. Chemom. 10, 47–61.

Bro, R., 1998. Multi-way Analysis in the Food Industry: Models, Algorithms and Applications. PhD Thesis. University of Amsterdam (NL) and Royal Veterinary and Agricultural University (DK).

Carroll, J.D., Chang, J., 1970. Analysis of individual differences in multidimensional scaling via an N-way generalization of '*Eckart–Young*' decomposition. Psychometrika 35, 283–319.

Casanova, R., Valdés-Sosa, P.A., García, F., Aubert, E., Riera, J., Korin, W., Lins, O., 2000. Frequency domain distributed inverse solutions. In: Aine, C.J., Okada, Y., Stroink, G., Swithenby, S.J., Wood, C.C. (Eds.), Biomag 96, Proceedings of the 10th International Conference on Biomagentism. Springer-Verlag, New York (ISBN:0387989153).

Chen, S., Donoho, D., Saunders, M., 2001. Atomic decomposition by basis pursuit. SIAM Rev. 43 (1), 129–159.

Churchland, P.S., Sejnowski, T.J., 1988. Perspectives on cognitive neuroscience. Science 242 (4879), 741–745.

Cohen, M.S., 1997. Parametric analysis of fMRI data using linear systems methods. NeuroImage 6 (2), 93–103.

Dahlhaus, R., 1997. Fitting time series models to non-stationary processes. Ann. Stat. 25, 1–37.

Davison, A.C., Hinkley, D.V., 1997. In: Gill, R., Ripley, B.D., Ross, S., Stein, M., Williams, D. (Eds.), Bootstrap Methods and Their Application. Cambridge Univ. Press, UK.

de Jong, S., Phatak, A., 1997. Partial least squares regression. Recent advances in total least squares techniques and errors—in variables modeling. In: Van Huffel, E. (Ed.), SIAM, Philadelphia.

Düzel, E., Habib, R., Schott, B., Schoenfeld, A., Lobaugh, N., McIntosh, A.R., Scholz, M., Heinze, H.J., 2003. A multivariate, spatiotemporal analysis of electromagnetic time–frequency data of recognition memory. NeuroImage 18, 185–197.

Estienne, F., Matthijs, N., Massart, D.L., Ricoux, P., Leibovici, D., 2001. Multi-way modelling of high-dimensionality electroencephalographic data. Chemom. Intell. Lab. Syst. 58, 59–71.

Field, A.S., Graupe, D., 1991. Topographic component (Parallel Factor) analysis of multichannel evoked potentials: practical issues in trilinear spatiotemporal decomposition. Brain Topogr. 3 (4), 407–423.

Friston, K.J., Frith, C.D., Fletcher, P., Liddle, P.F., Frackowiak, R.S.J., 1996. Functional topography: multidimensional scaling and functional connectivity in the brain. Cereb. Cortex 6, 156–164.

Fuchs, M., Wagner, M., Kohler, T., Wischman, H.A., 1999. Linear and nonlinear current density reconstructions. J. Clin. Neurophysiol. 16 (3), 267–295.

Galán, L., Biscay, R., Rodríguez, J.L., Pérez Abalo, M.C., Rodríguez, R., 1997. Testing topographic differences between event related brain potentials by using non-parametric combinations of permutation tests (published erratum appears in Electroencephalogr. Clin. Neurophysiol. 107(1998 Nov)(5): 380–381). Electroencephalogr. Clin. Neurophysiol. 102 (3), 240–247.

Goldman, R.I., Stern, J.M., Engel, J., Cohen, M.S., 2000. Acquiring simultaneous EEG and functional MRI. Clin. Neurophysiol. 111, 1974–1980.

Goldman, R.I., Stern, J.M., Engel, J., Cohen, M.S., 2002. Simultaneous EEG and fMRI of the alpha rhythm. NeuroReport 13 (18), 2487–2492.

Gonzalez-Andino, S.L., Blanke, O., Lantz, G., Thut, G., Grave de Peralta, R., 2001. The use of functional constraints for the neuroelectromagnetic inverse problem: alternatives and caveats. IJBEM 3, 1.

Harshman, R.A., 1970. Foundations of the PARAFAC procedure: models and conditions for an 'explanatory' multi-modal factor analysis. UCLA Work. Pap. Phon. 16, 1–84.

Horwitz, B., Poeppel, D., 2002. How can EEG/MEG and fMRI/PET data be combined? Hum. Brain Mapp. 17, 1–3.

Ioannides, A.A., 1999. Problems associated with the combination of MEG and fMRI data: theoretical basis and results in practice. In: Yoshimoto, T., Kotani, M., Kuriki, S., Karibe, H., Nakasato, N. (Eds.), Recent Advances in Biomagnetism. Tohoku University. Press, Sendai, pp. 133–136.

Kiers, H.A.L., 1991. Hierarchical relations among three-way methods. Psychometrica 56, 449–470.

Koenig, T., Martí-López, F., Valdés-Sosa, P.A., 2001. Topographic time–frequency decomposition of the EEG. NeuroImage 14, 383–390.

Kruskal, J.B., 1976. More factors than subjects, test and treatments: an indeterminacy theorem for canonical decomposition and individual differences scaling. Psychometrika, 41.

Kruskal, J.B., 1977. Three-way arrays: rank and uniqueness of trilinear decomposition with applications to arithmetic complexity and statistics. Linear Algebra Appl. 18, 95–138.

Lobaugh, N.J., West, R., McIntosh, A.R., 2001. Spatiotemporal analysis of experimental of experimental differences in event-related potential data with partial least squares. Psychophysiology 38, 517–530.

Martens, H., Naes, T., 1989. Multivariate Calibration. Wiley, Chichester.

McIntosh, A.R., Bookstein, F.L., Haxby, J.V., Grady, C.L., 1996. Spatial pattern analysis of functional brain images using Partial Least Square. NeuroImage 3, 143–157.

Miwakeichi, F., Martínez-Montes, E., Valdés-Sosa, P.A., Mizuhara, H., Nishiyama, N., Yamaguchi, Y., 2004. Decomposing EEG data into space–time–frequency components using parallel factor analysis. NeuroImage 22, 1035–1045.

Möcks, J., 1988a. Decomposing event-related potentials: a new topographic components model. Biol. Psychol. 26, 199–215.

Möcks, J., 1988b. Topographic components model for event-related potentials and some biophysical considerations. IEEE Trans. Biomed. Eng. 35, 482–484.

Pascual-Marqui, R.D., 1999. Review of methods for solving the EEG inverse problem. Int. J. Bioelectromagn. 1 (1), 75–86.

Pascual-Marqui, R.D., Valdés-Sosa, P.A., Alvarez, A., 1988. A parametric model for multichannel EEG spectra. Int. J. Neurosci. 40, 89–99.

Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18, 49–65.

Priestley, M.B., 1965. Evolutionary spectra and non-stationary processes. J. R. Stat. Soc., Ser. B Stat. Methodol. 27, 204–237.

Schneider, T., Neumaier, A., 2001. Algorithm 808: ARfit—A Matlab package for the estimation of parameter and eigenmodes of multivariate autoregressive models. ACM Trans. Math. Softw. 27 (1), 58–65.

Singh, M., Patel, P., Al-Dayeh, L., 1998. fMRI of brain activity during alpha rhythm. Int. Soc. Mag. Res. Med. 3, 1493.

Ståhle, L., 1989. Aspects of analysis of three-way data. Chemom. Intell. Lab. Syst. 7, 95–100.

Szava, S., Valdés-Sosa, P.A., Biscay, R., Galán, L., Bosch, J., Clark, I., Jímenez, J.C., 1994. High resolution quantitative EEG analysis. Brain Topogr. 6, 211–219.

Tallon-Baudry, C., Bertrand, O., 1999. Oscillatory gamma activity in humans and its role in object representation. Trends Cogn. Sci. 3 (4), 151–162.

Thomson, D.J., 1982. Spectrum estimation and harmonic analysis. Proc. IEEE 70, 1055–1096.

Trujillo, N.J., Martínez, E., Melie, L., Valdés, P.A., 2001. A Symmetrical Bayesian Model for fMRI and EEG/MEG Neuroimage Fusion. IJBEM 3, 1.

Valdés-Sosa, P.A., Bosch, J., Virués, T., Aubert, E., Fermín, E., González, E., 1998. EEG source frequency domain SPM. NeuroImage 7 (4), 636.