Blind Estimation of Channel Parameters and Source Components for EEG Signals: A Sparse Factorization Approach

Yuanqing Li, Andrzej Cichocki, and Shun-Ichi Amari, *Fellow, IEEE*

***Abstract*—In this paper, we use a two-stage sparse factorization approach for blindly estimating the channel parameters and then estimating source components for electroencephalogram (EEG) signals. EEG signals are assumed to be linear mixtures of source components, artifacts, etc. Therefore, a raw EEG data matrix can be factored into the product of two matrices, one of which represents the mixing matrix and the other the source component matrix. Furthermore, the components are sparse in the time-frequency domain, i.e., the factorization is a sparse factorization in the time frequency domain. It is a challenging task to estimate the mixing matrix. Our extensive analysis and computational results, which were based on many sets of EEG data, not only provide firm evidences supporting the above assumption, but also prompt us to propose a new algorithm for estimating the mixing matrix. After the mixing matrix is estimated, the source components are estimated in the time frequency domain using a linear programming method. In an example of the potential applications of our approach, we analyzed the EEG data that was obtained from a modified Sternberg memory experiment. Two almost uncorrelated components obtained by applying the sparse factorization method were selected for phase synchronization analysis. Several interesting findings were obtained, especially that memory-related synchronization and desynchronization appear in the alpha band, and that the strength of alpha band synchronization is related to memory performance.**

***Index Terms*—Electroencephalogram (EEG), linear mixture, linear programming, sparse factorization, synchronization, wavelet packets.**

I. Introduction

SPARSE factorization (also known as sparse representation or sparse coding) of signals has received tremendous attention in recent years [1]–[5] and many valuable theoretical results have been obtained [1]–[5]. Donoho and Elad [5] discussed optimal sparse representation in general (nonorthogonal) dictionaries via l^1 minimization. An interesting result that they obtained is that less than 50% of the concentration implies an equivalence between the l^0-norm solution and the l^1-norm solution. The sparse factorization approach can be used in blind source separation (BSS). In several recent studies, the mixing matrix and the sources were estimated using the maximum posterior approach, the maximum likelihood approach, and the expectation maximization algorithm, etc., [6]–[11]. A two-step approach is often used for BSS, in which the mixing matrix is estimated using the K-means or C-means clustering method, while the sources are estimated using a linear programming method [6]. Li and his colleagues discussed sparse representation and the applications of the two-step approach in BSS [12], [13]. The authors used a probabilistic approach and obtained the equivalence results of the l^0-norm solution and the l^1-norm solution. These results showed that if the sources are sufficiently sparse in the analyzed domain, they are more likely to be equal to the l^1-norm solution, which can be obtained using a linear programming method. However, precisely estimating the mixing matrix remains a fundamental problem in the two-step approach. Fuzzy C-means clustering and K-means clustering algorithms are locally convergent, and are sensitive to the lack of source sparseness; thus, they are not very effective in estimating the mixing matrix.

The electroencephalogram (EEG) signals are often assumed to be linear mixtures of electrical signals originating from multiple brain sources, artifacts, etc., which are considered as source components [14]–[17]. However, this assumption has raised two fundamental problems. The first is how to prove that this assumption is reasonable. Some researchers believe that it is obvious or even trivially true. However, one found that it is difficult to determine what the brain sources are and whether or not their locations are stable. Until now, to the best of our knowledge, there is no direct evidence that supports this assumption. The second problem is how to estimate the mixing process and brain sources if the linear mixture assumption is true. Independent component analysis (ICA), an approach commonly used in BSS [7]–[9], may be used to estimate the physiological brain sources, and in fact, has been often used to reveal the source components of EEG signals and has yielded many promising results [14]–[17]. However, because it works under the strong condition that all of the sources are mutually independent, which we do not think is reasonable, some estimated independent components might not be general brain sources.

This paper focuses on the application of sparse factorization in EEG data analysis. We also assume that EEG signals are linear mixtures of several source components, and that the

Manuscript received December 10, 2004; revised August 3, 2005. The work of Y. Li was supported by National Natural Science Foundation of China under Grant 60475004, the Program for New Century Excellent Talents in China University under Grant NCET-04-0816, Guangdong Province Science Foundation for Research Team Program under Grant 04205783.

Y. Li is with the Institute of Automation Science and Engineering, South China University of Technology, Guangzhou 510641, China (e-mail: yqli2@i2r.a-star.edu.sg).

A. Cichocki is with the Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Wako shi, Saitama 3510198, Japan (e-mail: cia@brain.riken.jp).

S.-I. Amari is with the Laboratory for Mathematical Neuroscience, RIKEN Brain Science Institute, Wako shi, Saitama 3510198, Japan (e-mail: amari@brain.riken.jp).

Digital Object Identifier 10.1109/TNN.2005.863424

number of components is greater than the number of electrodes. Under this assumption, the source components are sparse in the time-frequency domain. Theoretical and experimental data analysis not only propose a new algorithm for estimating the mixing matrix, but also lead to several evidences that support our assumption. Simulation experiment results based on artificial data show that our algorithm can better estimate the mixing matrix than the standard K-means clustering algorithm. After obtaining the mixing matrix, the source components can be estimated in the time frequency domain using a linear programming method. The source components in the time domain can be obtained by an inverse wavelet packets transformation. With the above mentioned two-step approach, an experimental EEG data matrix obtained from a subject performing a memory task is decomposed into a product of a mixing matrix and a source component matrix. A pair of almost uncorrelated source components is selected for phase synchronization analysis, leading to several interesting results. The phenomena of synchronization and desynchronization in the alpha band of the two components are observed. The average synchronization index for correct behavioral responses is higher than that for incorrect responses in most of the task time periods. This indicates that the strength of alpha rhythm synchronization is related to memory performance.

This paper is organized as follows. Section II presents a linear, overcomplete mixing model for EEG signals, which is a sparse factorization model of the EEG data matrix in the time frequency domain. A new algorithm for estimating the mixing matrix is proposed in Section III. Two examples and several evidences are given in Section IV to illustrate the algorithm and demonstrate the reasonability of the linear model representing the EEG signals. Section V discusses the estimation of source components. An application example about the phase synchronization analysis of the source components is presented in Section VI. Concluding remarks in Section VII summarize the advantages of the proposed approach.

II. Linear Mixture Model of EEG Signals

We first make the following assumption.

Assumption 1: EEG signals are linear time-invariant mixtures of source components such as brain sources, artifacts, etc. Note that if these signals are linearly and slowly time-varying mixtures, they also can be represented as linear time-invariant mixtures over a short time interval.

From Assumption 1, we have the following linear model:

$$\mathbf{X} = \mathbf{AS} + \mathbf{V} \tag{1}$$

where $\mathbf{X} = [\mathbf{x}(1), \ldots, \mathbf{x}(N)] = [x_i(j)]_{n\times N}$ is a known EEG data matrix of which each row is an EEG signal; N is the number of samples; $\mathbf{A} = [\mathbf{a}_1 \cdots \mathbf{a}_m]$ is a $n \times m$ unknown mixing constant matrix; $\mathbf{S} = [\mathbf{s}(1), \ldots, \mathbf{s}(N)] = [s_i(j)]_{m\times N}$ is an unknown source components matrix (,) of which the rows represent brain sources, artifacts, etc.; and $\mathbf{V} \in R^{n\times N}$ is an additive noise matrix. If n is not too large, which is 14 in examples of this study, we can assume that $m > n$, which implies that the model is overcomplete.

In the following theoretical analysis, we only consider the noise-free model corresponding to (1):

$$\mathbf{X} = \mathbf{AS}. \tag{2}$$

In several previous studies [14]–[18], linear model (1) or (2) was used. Parra and Sajda [19] provided converging evidence to support this assumption. However, there is no strong or direct evidence to show that the different sources in EEG are linearly superimposed, and there exist some fundamental and frequently asked questions, such as how to determine the brain sources and whether or not the number and locations of these sources are stable. Since it is difficult to answer these questions, it is difficult to justify that these different sources are linearly mixed on scalp electrodes. However, this does not prevent us from using a linear model in our EEG data analysis or from obtaining satisfying results.

Furthermore, as aforementioned, even if the linear mixture assumption is true, to date, there is no effective algorithm proposed to estimate the mixing matrix and source components. The ICA approach is often used to determine the mixing matrix and source components; however, all of the ICA algorithms are based on the assumption that all of the source components are mutually independent. It is reasonable to assume that the brain sources are independent of artifact sources, such as heartbeats and eye blinks. Therefore, the ICA approach is very effective in removing artifacts from EEG signals [16], [17]. However, it is not so reasonable to assume that all of the brain sources are mutually independent, and, hence, we do not believe that the ICA approach is ideal in determining the mixing matrix and the true brain sources.

In this paper, a sparse factorization approach is used for the determination of the mixing matrix and source components. This approach also provides evidence to support the Assumption 1. This approach consists of two main stages. In the first stage, a new algorithm is developed for estimating the mixing matrix, while the source components are estimated in the second stage.

The sparsity of source components plays a key role in this approach. EEG signals and the hidden source components are not sparse in the time domain; however, they are sparse in the time frequency domain. Therefore, our discussion in this study is mainly restricted in the time frequency domain. To obtain sufficiently sparse time frequency representations of EEG signals, we apply the wavelet packets transformation to (2) instead of a general wavelet transformation. Taking into account that the wavelet packets transformation is a linear transformation, we can have

$$\tilde{\mathbf{X}} = \mathbf{A}\tilde{\mathbf{S}} \tag{3}$$

where each row of $\tilde{\mathbf{X}}$ is a time-frequency representation of a corresponding EEG signal in $\mathbf{X}$ and each row of $\tilde{\mathbf{S}}$ is a time-frequency representation of a corresponding source component in $\mathbf{S}$.

It is well known that $\tilde{\mathbf{X}}$ in (3) should be sparse due to the wavelet packets transformation. If the EEG signals can be represented by a linear model (2), then $\tilde{\mathbf{S}}$ in (3) should also be very sparse (much sparser than $\tilde{\mathbf{X}}$). Thus, there should exist many

columns of $\tilde{\mathbf{S}}$ with only one nonzero entry. In a noisy environment, this means single entry dominant vectors. For instance, we can say that $\{\tilde{\mathbf{s}}(i_1),\ldots,\tilde{\mathbf{s}}(i_K)\}$ are K columns of $\tilde{\mathbf{S}}$, where only the first entry of each of these columns is nonzero; then, we have $\mathbf{A}\tilde{\mathbf{s}}(i_j)=\mathbf{a}_1\tilde{s}_1(i_j)$ and

$$\begin{aligned}[\tilde{\mathbf{X}}(i_1),\ldots,\tilde{\mathbf{X}}(i_K)] &= \mathbf{A}[\tilde{\mathbf{s}}(i_1),\ldots,\tilde{\mathbf{s}}(i_K)] \\ &= [\mathbf{a}_1\tilde{s}_1(i_1),\ldots,\mathbf{a}_1\tilde{s}_1(i_K)]. \end{aligned} \quad (4)$$

We now calculate a ratio matrix using the wavelet packets transformation coefficients

$$\tilde{\tilde{\mathbf{X}}} = \begin{bmatrix} \frac{\tilde{x}_1(1)}{\tilde{x}_q(1)} & \cdots & \frac{\tilde{x}_1(N)}{\tilde{x}_q(N)} \\ \vdots & \vdots & \vdots \\ \frac{\tilde{x}_n(1)}{\tilde{x}_q(1)} & \cdots & \frac{\tilde{x}_n(N)}{\tilde{x}_q(N)} \end{bmatrix}, \quad (5)$$

where $q\in\{1,\ldots,n\}$. In (5), we do not consider the case in which some $\tilde{x}_q(j)=0$. This case will be considered in the algorithm in Section III.

From (4), we can find a submatrix of $\tilde{\tilde{\mathbf{X}}}$ as follows:

$$\begin{bmatrix} \frac{\tilde{x}_1(i_1)}{\tilde{x}_q(i_1)} & \cdots & \frac{\tilde{x}_1(i_K)}{\tilde{x}_q(i_K)} \\ \vdots & \vdots & \vdots \\ \frac{\tilde{x}_n(i_1)}{\tilde{x}_q(i_1)} & \cdots & \frac{\tilde{x}_n(i_K)}{\tilde{x}_q(i_K)} \end{bmatrix} = \begin{bmatrix} \frac{a_{11}}{a_{q1}} & \cdots & \frac{a_{11}}{a_{q1}} \\ \vdots & \vdots & \vdots \\ \frac{a_{n1}}{a_{q1}} & \cdots & \frac{a_{n1}}{a_{q1}} \end{bmatrix}. \quad (6)$$

Furthermore, we have

$$\mathbf{a}_1 = a_{q1}\left[\text{mean}\left(\frac{\tilde{x}_1(i_j)}{\tilde{x}_q(i_j)}\right),\ldots,\text{mean}\left(\frac{\tilde{x}_n(i_j)}{\tilde{x}_q(i_j)}\right)\right]' \quad (7)$$

where mean $((\tilde{x}_l(i_j)/\tilde{x}_q(i_j)))=(1/K)\sum_{j=1}^{K}(\tilde{x}_l(i_j)/\tilde{x}_q(i_j))$.

Since the noise always exists in reality, we use the mean operation to estimate a column vector of the mixing matrix in (7) as well as in the algorithm in Section III.

Thus, if we determine a submatrix of $\tilde{\tilde{\mathbf{X}}}$, which has almost identical columns, we can estimate a column of the mixing matrix $\mathbf{A}$ up to a scale similarly as in (7).

If plotting all entries of the submatrix in (6) with respect to the column index i_j, we will obtain n horizontal lines. The key point in the algorithm in Section III is to detect the n horizontal lines in order to determine a submatrix of $\tilde{\tilde{\mathbf{X}}}$ with almost identical columns.

From the analysis above, we have the following conclusion.

Proposition 1: If EEG signals are linear mixtures of source components and if these source components are sufficiently sparse in the time frequency domain, then, the ratio matrix $\tilde{\tilde{\mathbf{X}}}$ in (5) contain several submatrices each of which has almost identical columns. The plot of all entries of such a submatrix with respect to their column indices can form n horizontal lines.

We analyzed many sets of raw EEG data from different experiments and different EEG machines. Our results support Proposition 1 (see Example 1 in Section IV). In Section IV, computational evidences based on raw EEG data will be presented to support Proposition 1 and Assumption 1.

Remark 1: One of the main tasks of this paper is to estimate the mixing matrix $\mathbf{A}$ and source component matrix $\mathbf{S}$ in (2) using the observed EEG data matrix $\mathbf{X}$ only. From the above analysis, we can see that it is possible to estimate mixing matrix $\mathbf{A}$ if matrix $\tilde{\mathbf{S}}$ in (3) is sufficiently sparse, or more precisely, if some column vectors of $\tilde{\mathbf{S}}$ are very sparse. Based on the estimated mixing matrix $\mathbf{A}$ and the sparsity of $\tilde{\mathbf{S}}$, $\tilde{\mathbf{S}}$ in (3), the time-frequency representation matrix of $\mathbf{S}$ can also be estimated using a standard linear programming method [12]. Here, source component matrix $\mathbf{S}$ can be obtained by using the inverse wavelet packets transformation on $\tilde{\mathbf{S}}$.

III. Algorithm for Estimating the Mixing Matrix

In this section, we propose a new algorithm for estimating mixing matrix $\mathbf{A}$. The key point this algorithm is to determine several submatrices of the ratio matrix $\tilde{\tilde{\mathbf{X}}}$ in (5) such that each has almost identical columns. Such a submatrix can be obtained by detecting n horizontal lines as described in the previous section (see Fig. 1 in Example 1 of Section IV). After determining a submatrix with almost identical columns, we can estimate a column vector of mixing matrix $\mathbf{A}$ up to a scale similarly as in (7).

Algorithm 1:

Step 1: Apply a wavelet packets transformation to every EEG signal represented by a row of matrix $\mathbf{X}\in R^{n\times N}$ and obtain a time-frequency representation matrix $\tilde{\mathbf{X}}$. The purpose of this step is to produce sparsification.

Step 2: Find a submatrix $\hat{\mathbf{X}}$ of $\tilde{\mathbf{X}}$ such that the norm of each column in it is greater than ξ_1, where ξ_1 is a prechosen positive constant. The purpose of this step is to reduce the computational burden of the estimation process and to remove the columns that are disproportionately influenced by noise.

Step 3: Let $n_1=1$ to n; then, do the followings (Loop 1 includes Steps 3.1 and 3.2.).

Step 3.1: Construct a ratio matrix $\hat{\tilde{\mathbf{X}}} = [(\hat{x}_k(i_j)/\hat{x}_{n_1}(i_j))]_{n\times K_1}$ using the entries of $\hat{\mathbf{X}}$, where K_1 is less than or equal to the column number of $\hat{\mathbf{X}}$. If the absolute value of an n_1-th row entry of $\hat{\mathbf{X}}$ is less than a preset positive constant ξ_2, we shall ignore the corresponding column of $\hat{\mathbf{X}}$ that contains the entry. This step is similar to the calculation in (5), thus, we use the same notation to represent the two ratio matrices.

Note: After Step 3.1, matrix $\hat{\tilde{\mathbf{X}}}$ contains several submatrices each of which has almost identical columns, i.e., its n rows can form n horizontal lines (see Fig. 1 in Example 1). The heights of these horizontal lines correspond to the column entries of the mixing matrix up to a scale.

The task of the following steps is to determine these submatrices with their rows representing horizontal lines.

Step 3.2: Let $n_2 = 1$ to n, $n_2 \neq n_1$ and do the followings (Loop 2 includes Steps 3.2.1, 3.2.2, and 3.2.3.).

Step 3.2.1: Find the minimum $\tilde{r}_{n_2}$ and maximum $\tilde{R}_{n_2}$ of $\hat{\tilde{\mathbf{x}}}_{n_2}$, which is the n_2-th row of $\hat{\tilde{\mathbf{X}}}$. Equally divide the interval $[\tilde{r}_{n_2}, \tilde{R}_{n_2}]$ into M_0 subintervals (bins), where M_0 is a preset positive integer. Divide matrix $\hat{\tilde{\mathbf{X}}}$ into M_0 submatrices, which are denoted by $\tilde{\hat{\tilde{\mathbf{X}}}}_1, \ldots, \tilde{\hat{\tilde{\mathbf{X}}}}_{M_0}$, such that all entries of the n_2-th row of $\tilde{\hat{\tilde{\mathbf{X}}}}_k$ are in the k-th subinterval, $k = 1, \ldots, M_0$.

Step 3.2.2: Let J_1 be a prechosen positive integer, and delete those submatrices of which the number of columns is less than J_1 from the submatrix set $\{\tilde{\hat{\tilde{\mathbf{X}}}}_k : k = 1, \ldots, M_0\}$. The remaining subset of submatrices is then denoted by $\{\tilde{\hat{\tilde{\mathbf{X}}}}_{j_k} : k = 1, \ldots, N_1\}$.

The objective of Steps 3.2.1 and 3.2.2 is to find the horizontal lines mentioned above by dividing one row's value interval of $\hat{\tilde{\mathbf{X}}}$ into many bins (see the first subplot of Fig. 7). However, these two steps are not sufficient to achieve this goal, and hence, we need to reduce the matrices obtained in Step 3.2.2.

Step 3.2.3: Let $n_3 = 1$ to n, $n_3 \neq n_1$ $n_3 \neq n_2$ and do the followings (Loop 3 includes Step a.

a: Let $k = 1$ to N_1 and do the followings (Loop 4 includes Steps a1, a2, and a3.).

a1: In Step 3.2.1, replace the n_2 and $\hat{\tilde{\mathbf{X}}}$ with n_3 and $\tilde{\hat{\tilde{\mathbf{X}}}}_{j_k}$ and then perform a step similar to Step 3.2.1. This leads to M_0 submatrices of $\tilde{\hat{\tilde{\mathbf{X}}}}_{j_k}$, which are denoted by $\tilde{\hat{\tilde{\mathbf{X}}}}_q^{(j_k)}$, $q = 1, \ldots, M_0$.

a2: For submatrix set $\left\{\tilde{\hat{\tilde{\mathbf{X}}}}_q^{(j_k)}, q = 1, \ldots, M_0\right\}$ and prefixed positive integer J_2, perform Step 3.2.2 and obtain a new set of submatrices. From this new submatrix set, choose a matrix, e.g., $\tilde{\hat{\tilde{\mathbf{X}}}}_p^{(j_k)}$ such that the sum of variances of its n rows is the smallest.

Steps a1 and a2, which are similar to Steps 3.2.1 and 3.2.2, can reduce the submatrices obtained in Step 3.2.2. Through these two steps, a submatrix $\tilde{\hat{\tilde{\mathbf{X}}}}_p^{(j_k)}$ of $\hat{\tilde{\mathbf{X}}}$ can be obtained, and its n rows can form n clear horizontal lines (see the second subplot of Fig. 7).

a3: Calculate the mean of all column vectors in matrix $\tilde{\hat{\tilde{\mathbf{X}}}}_p^{(j_k)}$ and normalize the averaged column vector to the unit norm. We can then obtain an estimated column vector of mixing matrix $\mathbf{A}$, which is denoted by $\mathbf{e}_i$.

Step 4: After performing the above four loops, we can obtain a set of estimated columns, denoted by $\mathbf{E} = [\mathbf{e}_1, \ldots, \mathbf{e}_{N_0}]$. Each column of $\mathbf{E}$ is an estimate of one column in mixing matrix $\mathbf{A}$. Since there are several loops above, a column in $\mathbf{A}$ may be estimated several times and all of the estimates are stored in $\mathbf{E}$. Therefore, $\mathbf{E}$ may have more columns than $\mathbf{A}$.

In this step, we must remove the duplication of column vectors in $\mathbf{E}$. If there is a set of several column vectors in $\mathbf{E}$ such that the angles between all pairs of these vectors are small (e.g., less than a predefined positive constant), then, we think that these vectors are redundant. We can calculate and, then, normalize their mean direction vector as a column estimation of the mixing matrix. Finally, the obtained matrix, denoted by $\bar{\mathbf{A}} \in R^{n \times m_0}$ $(m_0 > m)$, is taken as the estimate of the original mixing matrix $\mathbf{A}$. This step can also be carried out using the K-means clustering algorithm.

End of Algorithm 1.

In Algorithm 1, there exist five parameters which should be set in advance. In principle, their values depend on the data. ξ_1 and ξ_2 are related to the amplitude of the entries of the data matrix in the analyzed domain. Let Q_1 denote the maximum of the norms of all columns of the data matrix and Q_2 denote the maximum of the amplitude of the entries of the data matrix. We set the parameter ξ_1 to reduce noise and computation burden as stated in the algorithm. We can set ξ_1 to be a fraction of Q_1 (e.g., $0.3Q_1$). For implementing division operation, ξ_2 can be set as a fraction Q_2 (e.g., $0.1Q_2$). M_0 is the number of bins, which is set to be 100 in our simulations. J_1 and J_2 are the low bounds of the column number of the selected submatrices. Generally, $J_2 \leq J_1$. The general principle for setting M_0, J_1, and J_2 is to make sure that these horizontal lines can be observed clearly. In our simulations, we take the positive integers J_1 and J_2 in $[50,100]$.

In Appendix, we present an example based on artificially mixed data to illustrate Algorithm 1. In that example, the mixing matrix can be successfully estimated using our algorithm. When compared to the standard K-means clustering algorithm, Algorithm 1 is not as sensitive to the lack of $\tilde{\mathbf{S}}$ sparseness. The example in Appendix indicates that, although only 8% of $\tilde{\mathbf{S}}$ columns are very sparse (i.e., have only one nonzero entry), the mixing matrix can be precisely estimated using Algorithm 1. The standard K-means clustering algorithm does not work well

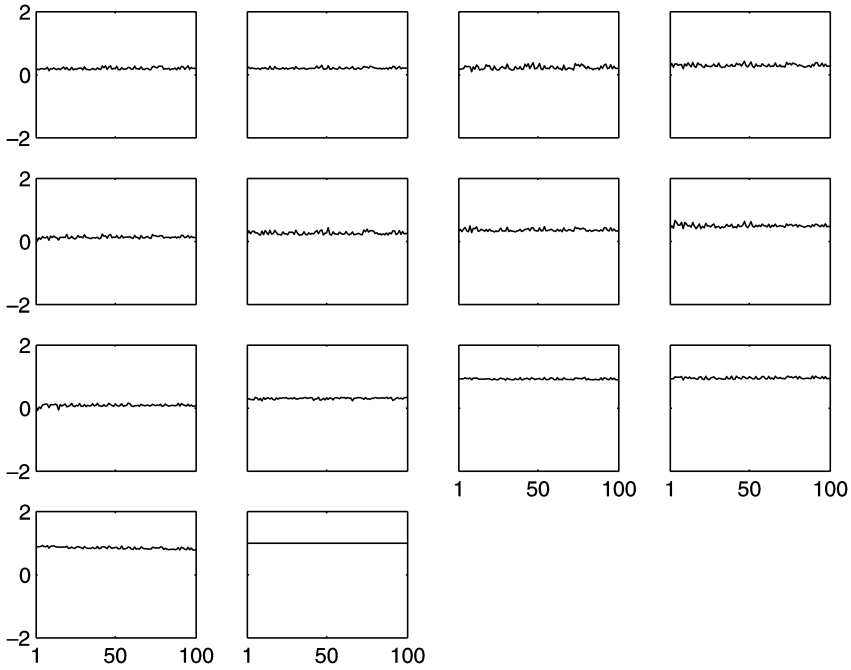

Fig. 1. Fourteen ratio curves calculated by (6) in Example 1.

in this case. Therefore, our algorithm can obviously better estimate the mixing matrix than the standard K-means clustering algorithm. More detailed discussions on this algorithm can be seen in [37].

In Sections IV–VI, we will use Algorithm 1 to estimate channel parameters (mixing matrix) for raw EEG data.

IV. Examples and Evidences

In this section, we present two examples using Algorithm 1 based on raw EEG data matrices and provide four evidences showing that Assumption 1, i.e., the linear model assumption in Section II, is reasonable for EEG signals.

Example 1: Consider a raw EEG data matrix $\mathbf{X} \in R^{14\times 20\,000}$ of which each row is a channel EEG signal recorded from an electrode. Applying the discrete wavelet packets transformation with seven levels to each EEG signal of $\mathbf{X}$, we obtain the time frequency representation matrix $\tilde{\mathbf{X}} = [\tilde{\mathbf{X}}_1, \ldots, \tilde{\mathbf{X}}_7]$, where the submatrix $\tilde{\mathbf{X}}_i$ corresponds to the i-th node of the wavelet packets tree $i = 1, \ldots, 7$.

In this example, we first use Algorithm 1 to determine a submatrix with almost identical columns. If we obtain such a submatrix and plot all entries of the submatrix with respect to their column indices, then we can have 14 horizontal lines.

Because of the noise, we only consider those vectors of $\tilde{\mathbf{X}}$ with relatively large norms. We set $\xi_1 = (2M/3)$, where M is the norm maximum of all columns of $\tilde{\mathbf{X}}$. Removing all columns with 2-norms less than ξ_1, we obtain a matrix denoted as $\hat{\mathbf{X}}$ that is composed of all of the left columns.

For $n_1 = 14$, we first find the maximum of absolute values of all entries in the n_1-th row of $\hat{\mathbf{X}}$. Setting $\xi_2 = (M_{14}/5)$, we construct a ratio matrix $\hat{\tilde{\mathbf{X}}}$ according to Step 3.1 in Algorithm 1.

After setting $M_0 = 100$, $J_1 = J_2 = 80$, we continue to perform Step 3.2 to determine several submatrices of $\hat{\tilde{\mathbf{X}}}$ each of which has almost identical columns. One of the several submatrices is illustrated in Fig. 1.

There are 14 subplots in Fig. 1. The curve of the i-th subplot represents all entries in the i-th row of the submatrix with respect to their column indices. The 14 curves can be seen as horizontal lines.

The 14 subplots in Fig. 2 show all entries in the 14 rows of a submatrix which is composed by the first 300 columns of ratio matrix $\hat{\tilde{\mathbf{X}}}$ above. Comparing these curves in Figs. 1 and 2, we can say that no horizontal lines appear in the first 13 subplots.

Remarks 2: 1) The ratios shown in Fig. 1 are not strict constants. This is because a) noise still exists; and b) any sufficiently sparse column vector of matrix $\tilde{\mathbf{S}}$ in (3) is a single entry dominant vector of which only one entry has a large absolute value and the others are close to but not equal to zero. 2) When compared with the ratios shown in Fig. 2, we can conclude that the ratios in each subplot of Fig. 1 are close to a constant. That is, the obtained submatrix has almost identical columns.

The mean column vector of a submatrix with almost identical columns, determined above, is followed by normalization and is an estimated column vector of the mixing matrix up to a scale.

We used Algorithm 1 to estimate the mixing matrix and obtained more than 3500 estimated column vectors in this example. Here, $\mathbf{E}$ denotes the matrix containing these column vectors. Removing the redundant columns in $\mathbf{E}$, we can obtain the estimated mixing matrix $\bar{\mathbf{A}} \in R^{14\times 20}$, which is omitted here due to limited page space.

Example 1 provides us with two evidences supporting the assumption given in the previous section.

Evidence 1: The result of experimental data analysis in Example 1 provides firm evidence supporting Proposition 1. Since Proposition 1 is directly derived from Assumption 1, the result obtained in Example 1 also provides firm evidence supporting Assumption 1.

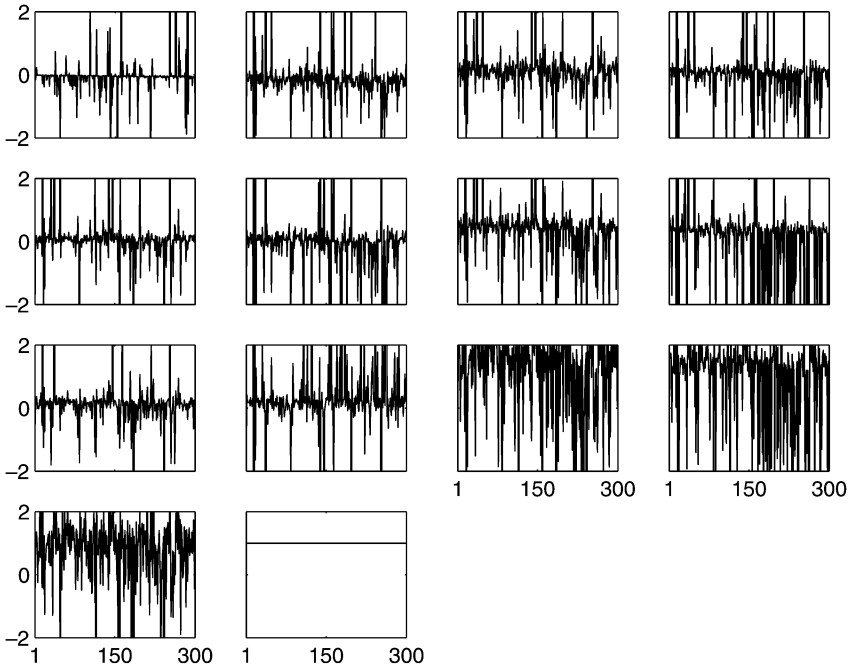

Fig. 2. Ratios calculated using the first 300 columns of $\hat{\mathbf{X}}$ in example 1.

Evidence 2: If Assumption 1 holds, then the mixing matrix should be nonnegative. In Example 1, we found that the estimated mixing matrix is a nonnegative matrix, which is not presented here because of limited page space. This result has also been confirmed in our other EEG data analysis examples.

Using the raw EEG data matrix in Example 1, we construct a data matrix as follows:

$$\mathbf{Y} = \mathbf{C}\mathbf{X} \tag{8}$$

where $\mathbf{C} \in R^{14\times 14}$ is a constant matrix representing a linear transformation.

After applying Algorithm 1 to the constructed data matrices $\mathbf{Y}$, we have Evidence 3, which follows.

Evidence 3: Applying Algorithm 1 to the constructed data matrices $\mathbf{Y}$, which can be represented by a linear model ($\mathbf{Y} = \mathbf{CAS}$) like (1) according to Assumption 1, we find that the conclusions in Proposition 2 still hold. Furthermore, the estimated mixing matrix is nonnegative if matrix $\mathbf{C}$ in (8) is nonnegative, and it is not nonnegative if matrix $\mathbf{C}$ in (8) is not nonnegative.

Example 2: Consider a combined EEG data matrix

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}$$

where $\mathbf{X}_1, \mathbf{X}_2 \in R^{n_1\times N}$ are two different raw EEG data matrices from different subjects, even from different EEG machines. This means that $\mathbf{X}_1$ and $\mathbf{X}_2$ have different source components.

Suppose that the source numbers are m_1 and m_2 for $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively. According to Assumption 1, we have

$$\mathbf{X}_1 = \mathbf{A}_1\mathbf{S}_1 \tag{9}$$

$$\mathbf{X}_2 = \mathbf{A}_2\mathbf{S}_2 \tag{10}$$

where $\mathbf{A}_1 \in R^{n_1\times m_1}$ and $\mathbf{A}_2 \in R^{n_1\times m_2}$ are two different mixing matrices, and $\mathbf{S}_1 \in R^{m_1\times N}$ and $\mathbf{S}_2 \in R^{m_2\times N}$ are two different source component matrices for $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively.

Combining the above two equations, we have

$$\mathbf{X} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix} \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_2 \end{bmatrix} \tag{11}$$

where the two $\mathbf{0}$s are two zero matrices with consistent dimensions.

If Assumption 1 in Section II holds for the EEG signals, an estimated column vector of the mixing matrix in (11) should have the form $[*, \mathbf{0}]'$ or $[\mathbf{0}, *]'$, where $*$ represents a n_1 dimensional nonzero row vector, and $\mathbf{0}$ is a n_1 dimensional zero row vector.

We next use two sets of raw EEG data matrices to estimate the mixing matrix in (11) and demonstrate the previously mentioned deduction.

1) $\mathbf{X}_1 \in R^{7\times 20\,000}$ was recorded in an experiment carried out in China (sampling rate: 1000 Hz), and $\mathbf{X}_2 \in R^{7\times 20\,000}$ was recorded in an experiment carried out in Russia (sampling rate: 256 Hz).

   Using the combined EEG data matrix $\mathbf{X} = [\mathbf{X}_1', \mathbf{X}_2']'$ and Algorithm 1, we can estimate many column vectors of the mixing matrix in (11). All of the estimated vectors have either the form $[*, \mathbf{0}]'$ or the form $[\mathbf{0}, *]'$. The two estimated column vectors are: [−0.0067, −0.0022, −0.0088, −0.0069, −0.0090, 0.0003, 0.0026, 0.2144, 0.3040, 0.2229, 0.4949, 0.1652, 0.7035, 0.2111]′ and [0.5101, 0.3873, 0.5349, 0.3614, 0.2834, 0.1564, 0.2287, −0.0537, −0.0681, 0.0622, −0.0143, −0.0375, −0.0242, −0.0492]′.

2) $\mathbf{X}_1, \mathbf{X}_2 \in R^{7\times 20\,000}$ are two raw EEG data matrices from the same EEG machine but from two different subjects.

Using the combined EEG data matrix $\mathbf{X}$ and Algorithm 1 to estimate the column vectors of the mixing matrix in (11), we obtain a similar result as in 1), and the two estimated columns are: [0.0455, 0.0312, 0.0246, 0.0082, 0.0045, 0.0336, 0.0324, 0.3928, 0.3562, 0.3678, 0.4021, 0.3493, 0.4032, 0.3625]′ and [0.4928, 0.4083, 0.3939, 0.2629, 0.2606, 0.3856, 0.3864, −0.0037, −0.0016, −0.0079, 0.0057, −0.0056, −0.0019, −0.0038]′.

From the analysis and calculations above, we have Evidence 4.

Evidence 4: A combined EEG data matrix, which is composed by two different EEG data matrices, as shown in Example 2, is used to estimate the column vectors of the mixing matrix. If the EEG signals can be represented by a linear mixing model (1), the estimated column vectors of the mixing matrix in (11) should have either the form $[*, \mathbf{0}]'$ or the form $[\mathbf{0}, *]'$. This has been demonstrated in calculation with real EEG data.

V. Linear Programming Algorithm for Estimating EEG Source Components

After obtaining the mixing matrix, the next step is to estimate the source component matrix, which is the main task of this section.

Since the estimated mixing matrix $\bar{\mathbf{A}}$ is not a square matrix, which has its column number being larger than its row number, the source component matrix cannot be obtained by solving the system of linear equations. In this study, we first estimate the time frequency representation matrix $\tilde{\mathbf{S}}$ in (3) by solving an optimization problem. We can then calculate the source component matrix $\mathbf{S}$ using an inverse wavelet packets transformation. For the estimated mixing matrix $\bar{\mathbf{A}}$ in (1), the time frequency representation matrix $\tilde{\mathbf{S}}$ of the source component matrix can be found by solving the following optimization problem:

$$\begin{aligned} \min \quad & \sum_{i=1}^{m}\sum_{j=1}^{N}|\tilde{s}_i(j)| \\ \text{subject to} \quad & \bar{\mathbf{A}}\tilde{\mathbf{S}} = \tilde{\mathbf{X}}. \end{aligned} \tag{12}$$

It is not difficult to prove that the linear programming problem (12) is equivalent to the following set of N smaller scale linear programming problems:

$$\begin{aligned} \min \quad & \sum_{i=1}^{m}|s_i(j)| \\ \text{subject to} \quad & \bar{\mathbf{A}}\tilde{\mathbf{s}}(j) = \tilde{\mathbf{x}}(j), \qquad j = 1, \ldots, N. \end{aligned} \tag{13}$$

By setting $\tilde{\mathbf{S}} = \tilde{\mathbf{U}} - \tilde{\mathbf{V}}$, where $\tilde{\mathbf{U}} = [\tilde{u}_i(j)]_{m\times N} \geq 0$ and $\tilde{\mathbf{V}} = [\tilde{v}_i(j)]_{m\times N} \geq 0$, (13) can be converted into the following standard linear programming problems with nonnegative constraints:

$$\begin{aligned} \min \quad & \sum_{i=1}^{m}(\tilde{u}_i(j) + \tilde{v}_i(j)) \\ \text{subject to} \quad & [\bar{\mathbf{A}}, -\bar{\mathbf{A}}][\tilde{\mathbf{u}}^T(j), \tilde{\mathbf{v}}^T(j)]^T = \tilde{\mathbf{x}}(j), \\ & \tilde{\mathbf{u}}(j) \geq 0, \qquad \tilde{\mathbf{v}}(j) \geq 0 \end{aligned} \tag{14}$$

where $j = 1, \ldots, N$.

After estimating $\tilde{\mathbf{S}}$, $\mathbf{S}$ can be obtained by an inverse wavelet packets transformation corresponding to the wavelet packets transformation used in (3). Combining the discussion of this section and the previous sections, we have the following algorithm for estimating the source components.

Algorithm 2:

Step 1. Transform the n raw EEG signals (n rows of $\mathbf{X}$) into n time-frequency signals by a wavelet packets transformation and obtain a time frequency matrix $\tilde{\mathbf{X}}$.

Step 2. Estimate the mixing matrix $\bar{\mathbf{A}} \in R^{n\times m'}$ using Algorithm 1.

Step 3. Using the estimated mixing matrix $\bar{\mathbf{A}}$ and matrix $\tilde{\mathbf{X}}$, estimate the time frequency representation matrix $\tilde{\mathbf{S}}$ in (3) by solving the set of linear programming problems (14) that are equivalent to (12).

Step 4. Reconstruct source component matrix $\mathbf{S}$ in (2) by using the inverse wavelet packets transformation.

End.

Remark 3: Using the above algorithm, we blindly, but successfully, recovered speech sources from several mixtures in an overcomplete case (eight sources, four sensors) [12].

Here, a recoverability question arises: Why can $\tilde{\mathbf{S}}$ be estimated by solving the linear programming problems (12) or (14)? In other words, why is the solution of (12) equal to the true solution $\tilde{\mathbf{S}}$ in (3)? We can find the answer in [12] (Theorem 6), which has presented the recoverability analysis results on BSS based on sparse representation.

Note that under the Assumption 1, the condition our algorithms work well is that the component matrix in (1) is sufficiently sparse in the analyzed domain. Thus, in Step 1 of Algorithms 1 and 2, we apply a wavelet packets transformation (instead of wavelet transformation) to EEG signals to produce sparsification. Obviously, more levels of used wavelet packets transformation there are, the higher sparsity of data is. However, large number of levels of wavelet packets transformation imply a heavy computation burden. In our simulations for EEG data analysis as well as blind separation of speech signals, we often use 7-level wavelet transformation.

VI. Phase Synchronization Analysis Based on Source Components

We now present an example of analyzing real EEG data as an application of our proposed approach. Three healthy male adults performed modified Sternberg memory tasks. The subjects were instructed to memorize three numbers successively presented at random positions on a computer monitor. After a 2.5 second pause, a warning signal was presented, and a "test number" was given on the monitor. If the test number was one of the memorized numbers, the subject had to press a button. This study-test cycle, which included a rest period, was repeated 160 times (about 24 min.).

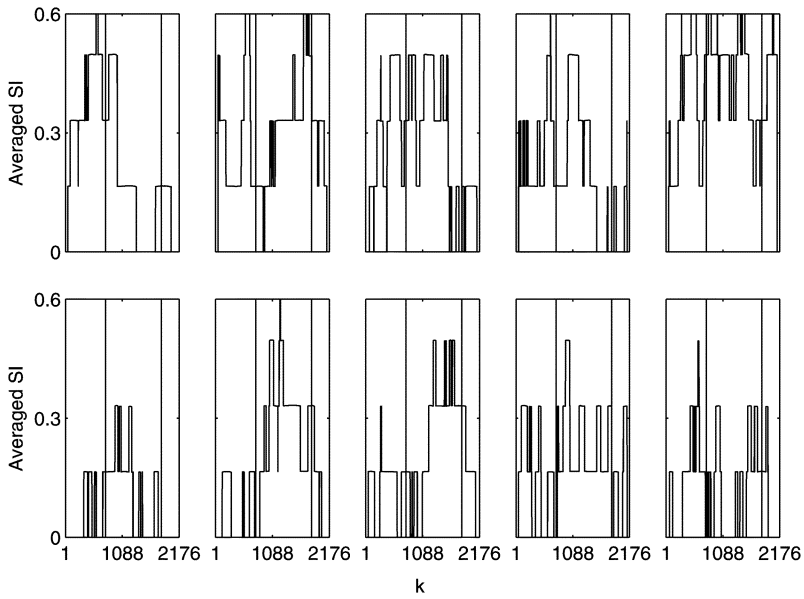

Fig. 3. Time course of EEG synchrony during single trials. First row: Mean synchronization index averaged across frequencies in the 8 Hz–13 Hz range for five single trials with correct responses. Second row: Mean synchronization index averaged across frequencies in the 8 Hz–13 Hz range for five single trials with incorrect responses. In each subplot, the first vertical line indicates the beginning of number presentations to be memorized in the Sternberg task, and the second vertical line indicates the end of test number presentation.

The EEG was recorded from the scalp by 16 electrodes according to the international 10–20 system (sampling frequency: 256 Hz, analog filtering frequency: 0 Hz–30 Hz). Only the EEG signals of the first 14 electrodes were used in our analysis.

Here, we only describe the analysis results of the first subject's data in detail and simply show the final results of the other two subjects.

A few trials seriously affected with artifacts (eye blinks and muscle noise) were rejected by visual inspection of the EEG record, and a data set comprised of 20 trials with correct responses and 20 trials with incorrect responses was selected for analysis (one trial = 2176 digitized voltage samples, which represented 8.5 s of the EEG). We, thus, obtained a $14 \times 87\,040$ dimensional data matrix, which is denoted by $\mathbf{X}$. Based on Algorithm 1, we also had a 14×20 dimensional mixing matrix denoted as $\bar{\mathbf{A}}$. Using Algorithm 2, we decomposed the EEG signal $\mathbf{X}$ into 20 source components, and denoted it as the $20 \times 87\,040$ dimensional components matrix $\mathbf{S}$. This contained 20 trials of correct responses and 20 trials of incorrect responses, respectively.

We first calculated the correlation coefficient matrices of $\mathbf{X}$ and $\mathbf{S}$, which were denoted by $\mathbf{R}^x = \left[r^x_{i,j}\right]_{14\times14}$ and $\mathbf{R}^s = \left[r^s_{i,j}\right]_{20\times20}$, respectively. We found that $r^x_{i,j} \in (0.2003, 1)$ (the median of $\left|r^x_{i,j}\right|$ is 0.5845). For the component cases, many component pairs had small correlation coefficients, such as $r^s_{2,12} = 0.0025$ and $r^s_{2,17} = 0.0167$, and the median of $\left|r^s_{i,j}\right|$ was 0.2367. Furthermore, we found that the higher order correlation coefficients of these pairs were also very small. It should be emphasized here that, although the independence principle and decorrelation principle were not used, many pairs of components were almost uncorrelated.

According to modern brain theories, the synchronization dynamics of rhythmic activity in distinct neural networks plays a very important role in the interactions among them [20], [21]. Research suggests that patterns of high coherence between EEG signals recorded at different scalp sites have functional significance and are correlated with different kinds cognitive information processing, including memory, language, concept retrieval, etc.,[22]. In relation to memory processes, studies have reported an increase of synchronization among brain regions involved in the respective task [23]. Allefeld and Kurths [24] proposed a method for multivariate phase synchronization analysis of EEG data, and their calculated synchronization indices indicated two distinct increases in synchronization related to stimulus presentation. In order to see the importance of synchronization and desynchronization of alpha rhythm in memory information processing [25], we mainly focus our attention on the analysis of alpha rhythm phase synchronization between two source components.

We selected a pair of two, almost uncorrelated, components $(\mathbf{s}_2, \mathbf{s}_{17})$ ($r^s_{2,17} = 0.0167$, fourth correlation coefficient -0.0032) and filtered them offline in a 1–70 Hz range for phase synchronization analysis. [26] Quiroga *et al.* analyzed several different synchronization indices and showed their consistency to a certain extent. In this study, we used the phase synchronization index described in [28], [29] in our analysis. The synchronization index was defined by $SI(f,t) = \max(\mathrm{SPLV(f,t)} - \mathrm{Ssur}, 0)$, where $\mathrm{SPLV(f,t)}$ was a single-trial phase-locking value at frequency f and time t; it was smoothed using a window with a length of 99. $S\mathrm{sur}$ in the 0.97 distribution percentile of 200 surrogates. The 200 pairs of surrogate data are Gaussian distributed.

Fig. 3 displays the analysis results of phase synchrony. The subplots of the first row present the mean synchronization index values, which were averaged across frequencies in the 8–13 Hz range (alpha band) for five single trials with correct responses. The subplots of the second row present the mean synchronization index values, which were also averaged across frequencies

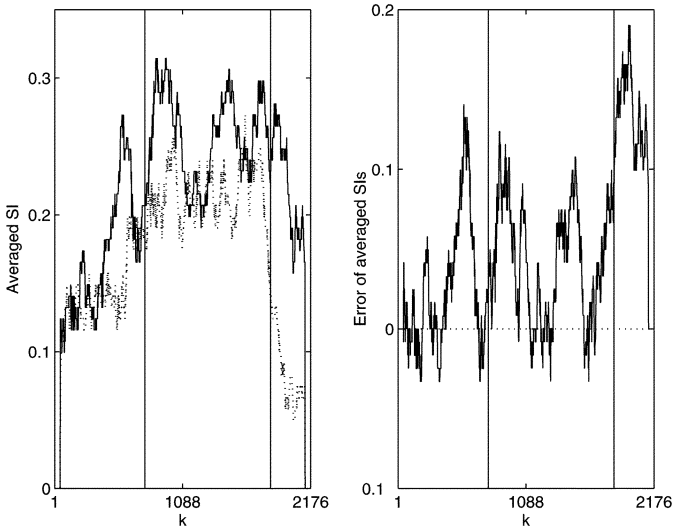

Fig. 4. (Left) Time course of EEG synchrony, averaged across trials. Same subject as in previous figure. Curves show mean values of the synchronization indices averaged in the 8–13 Hz range across 20 trials. The bold curves are trials with correct responses, and the dotted curves are trials with incorrect responses. Solid vertical lines are the same as in the previous figure. (Right) Subtraction of the averaged synchronization index for incorrect responses from that for correct responses (those in the left subplot).

in the 8–13 Hz range for five single trials with incorrect responses. In each subplot, the first vertical line refers to the beginning of the number presentation to be memorized, and the second line refers to the end of the test number presentation.

During a single trial, there exists one or several time intervals in which the synchrony is very strong, and, conversely, there also exists one or several time intervals in which the synchrony is very weak. These phenomena are called synchronization and desynchronization. Intuitively, we can also see that the synchronization indices for most trials with correct responses were higher than those for most trials with incorrect responses. Although only ten trials of the analyzed 40 trials were presented due to limited page space, the other 32 trials showed similar characteristics.

In the first subplot of Fig. 4, two averaged synchronization index curves are displayed. The two curves were obtained by averaging synchronization index SI in the 8–13 Hz range and across 20 trials, separately for correct and incorrect responses. The bold curve represents trials with correct responses and the dotted curve represents trials with incorrect responses. The second subplot of Fig. 4 shows the difference of the two averaged synchronization indices shown in the first subplot (averaged synchronization index for correct responses minus the averaged synchronization index for incorrect responses). From Fig. 4, we can see that during most of the time periods, especially at the beginning of the trial (preparation or attention period), the averaged synchronization index for the correct responses is greater than that for the incorrect responses. The higher synchronization for correct responses could be related to the higher integration of brain systems that is required for effective information processing.

We next analyzed the EEG data from the other two subjects. Based on the raw EEG data matrix $\mathbf{X}$, we also calculated the source component matrix $\mathbf{S}$ using the proposed approach. Two

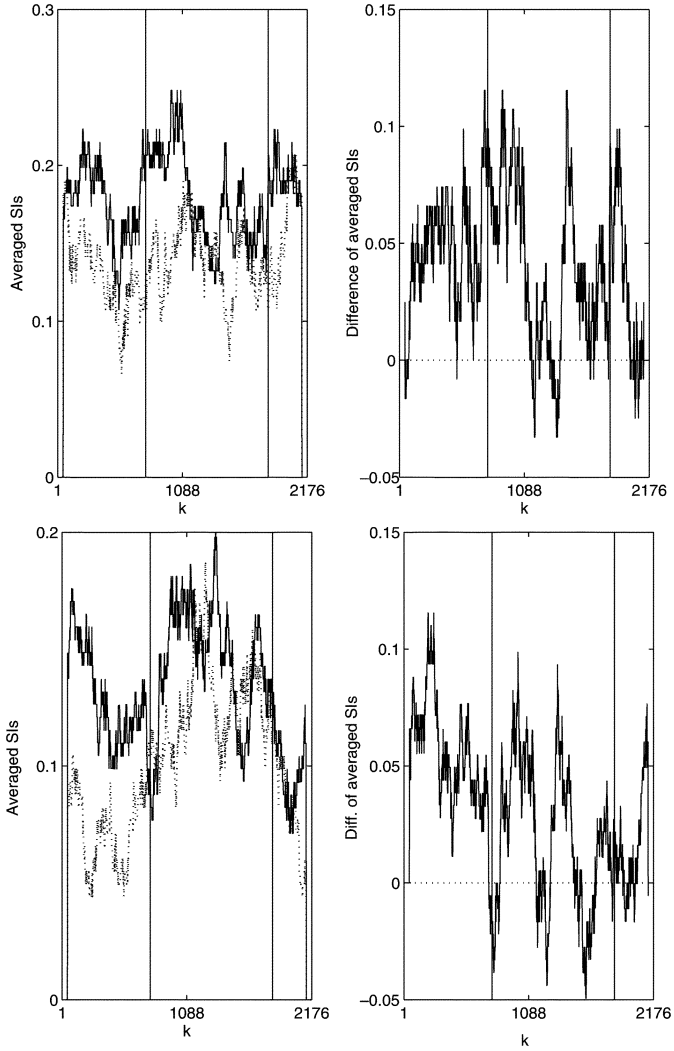

Fig. 5. Two subplots in the first row show the synchronization analysis results for the second subject. The two subplots in the second row are for the third subject. The caption for each row is similar to that in Fig. 4.

almost uncorrelated source components were chosen for phase synchronization analysis. Results showed the phenomena of synchronization and desynchronization similar to that in Fig. 3 for every trial. Due to the page space limit, we only give here the averaged synchronization indices. The subplots in Fig. 5 present our analysis results for the second and the third subjects. The first row: First subplot displays two averaged synchronization index curves, which were obtained by averaging synchronization index SI in the 8 Hz–13 Hz range and across 20 trials, separately for correct and incorrect responses. The bold curve refers to trials with correct responses, and the dotted curve refers to trials with incorrect responses. The second subplot shows the difference of the two averaged synchronization indices displayed in the first subplot. The second row: Two subplots are for the third subject, in which the first shows two averaged synchronization index curves for correct responses (black) and incorrect responses (dotted), respectively, while the second shows their difference.

Discussions and Conclusions:

1) To the best of our knowledge, almost all related references analyzed the synchronization between two of

EEG channels (or electrodes). A substantial part of synchronization between two EEG channels can be explained by volume conduction effects. In this study, we use the sparse decomposition approach to determine the source components. Two almost uncorrelated components were selected for phase synchronization analysis. One objective of this approach is to exclude the volume conduction influence. Here, we conclude that the phenomena of synchronization and desynchronization is not from volume conduction based on the following two reasons. First, the two analyzed components are almost uncorrelated, and their second and even higher order correlation coefficients are negligible. Second, if the observed synchronization is from volume conduction, no clear difference should exist between the averaged synchronization index for the correct responses and that for the incorrect responses.

2) During the attention period, alpha rhythm synchronization appeared for the three subjects, and the synchronization index for correct responses was stronger than that for incorrect responses. This indicates that synchronization strength is related to memory performance. We also found synchronization in the period of encoding and retrieval. This is consistent with observations in [30] that alpha rhythm synchronization appeared in retention time. Furthermore, we observed desynchronization between the attention and encoding periods. These results are in line with several previously obtained results. For instance, Klimesch argued that better memory performance is associated with a higher baseline alpha (before memory retrieval) and a greater alpha desynchronization [25]. Hogan *et al.* confirmed these [31].

Our findings are also consistent with the results in several previous studies dealing with modified Sternberg memory experiments [30]–[32]. In [30], phase-coupling of theta-gamma EEG rhythms during short-term memory processing was analyzed by bispectral analysis. [32] discussed event-related desynchronization (ERD)/event-related synchronization (ERS) during an auditory memory task. ERS was generated during the presentation of the memory set (encoding), and an ERD was generated in all of the subject groups during stimulus matching, which is the memory scanning phase (retrieval). This finding was correlated with earlier observations concerning alpha activity that alpha suppression was a correlate of cognitive processes in association with auditory information processing, such as short-term memory retrieval [33]. In [31], memory-related EEG power and coherence on temporal and central recording sites in patients with early Alzheimer's disease and normal controls was examined. It was found that the synchronization strength of the alpha band was related to memory performance.

In this section, EEG data analysis of the three subjects exhibited consistent results. Several pairs of almost uncorrelated components can always be found, which show memory-related synchronization and desynchronization. Our future investigation will focus on some other characteristics of the source components, such as localization, stability, etc. Additionally, it would be useful to investigate the phase-coupling of different frequency bands for different source components, which could be interpreted as an EEG aspect of the functional linking among the different brain areas. This can include the prefrontal areas and the G.cinguli (as part of the limbic system), which are both extremely important for memory functions [30].

In EEG data analysis, one often uses the time-frequency analysis, which is also a kind of sparse factorization under a given dictionary. For instance, under the given dictionary, some studies have discussed the time-frequency parameterization of EEG epileptic spikes [35], [36]. From Algorithms 1 and 2, and the data analysis in this section, we can see that there are three main stages in our approach: 1) time-frequency factorization (wavelet packets transformation) to produce sparsification of data; 2) sparse factorization in the time frequency domain for estimating the mixing matrix and components; and 3) synchronization analysis based on sparse components. Thus, our approach is different from the general sparse time-frequency decomposition under a given dictionary.

VII. Concluding Remarks

In this study, based on a linear mixture model of EEG signals, a two-step sparse factorization approach was employed instead of an ICA approach, to estimate the mixing matrix and the source components. In this approach, the first step is to estimate the mixing matrix, and the second step is to estimate the source component matrix. One of main tasks in this paper is to develop a new algorithm for estimating the mixing matrix.

Fundamentally, EEG signals are assumed to be linear overcomplete mixtures of source components. That is, the raw EEG data matrix can be factorized into a product of the mixing matrix and the component matrix. This factorization is a sparse factorization in the time frequency domain. By constructing a ratio matrix using the wavelet packets transformation coefficients of EEG signals, we conclude that the ratio matrix has several submatrices each of which has almost identical columns. This conclusion comes from the linear model assumption and sparseness of source components. Our extensive computational results are based on many sets of EEG data from different experiments, different subjects, and different EEG machines, and strongly support this conclusion. Furthermore, a new algorithm has been proposed for estimating the mixing matrix based on the conclusion. Two examples based on raw EEG data matrices are given not only to illustrate the algorithm, but also to provide firm evidences supporting our assumption for EEG signals.

After estimating the mixing matrix, the sparse wavelet packets transformation coefficients of the source components are calculated using a linear programming method. We can thus obtain a sparse decomposition of the EEG time frequency representation matrix. The source components in the time domain are then reconstructed by the corresponding inverse wavelet packets transformation. This construction employs sparse representation and can be used in blind source separation, especially when the source number is unknown, fewer sensors exist than sources, and the sources are not completely independent [12].

Finally, an example of phase synchrony analysis in real EEG data is given to support the validity and performance of

the proposed approach. A pair of almost uncorrelated source components was selected for phase synchrony analysis. The phenomena of synchronization and desynchronization are observed. The averaged synchronization index in the alpha band for correct responses is higher than that for incorrect responses in most of the time periods, especially in the beginning of the trial (preparation periods). Since the analyzed components are almost uncorrelated, we can conclude that it is not the volume conduction that evokes the synchronization.

As we know, many promising results for EEG data analysis have been obtained using the ICA approach, especially with regards to removing artifacts, such as eye blinks, eye movements, heartbeats, etc. When compared with ICA, the sparse factorization approach has two important advantages. First, sources do not necessarily have to be mutually independent, and can even be nonstationary. Second, the number of sources can be greater than the number of sensors and it can even be unknown. We believe that sparse factorization is an alternative to ICA and very promising approach for analyzing EEG.

APPENDIX
EXAMPLE OF ESTIMATING THE MIXING MATRIX BASED ON ARTIFICIALLY MIXED DATA

In this example, the source matrix $\mathbf{S} \in R^{4\times 10\,000}$ is drawn from a uniform distribution in $[-5, 5]$, in which there are 800 columns with only one nonzero entry. Among the 800 columns, 200 have nonzero first entries, 200 have nonzero second entries, etc. The mixing matrix $\mathbf{A} \in R^{3\times 4}$ is taken randomly followed by normalization as follows:

$$\mathbf{A} = \begin{bmatrix} 0.8412 & -0.0298 & -0.1735 & 0.7240 \\ -0.5025 & 0.8305 & 0.3621 & 0.4088 \\ 0.1997 & 0.5563 & 0.9158 & -0.5556 \end{bmatrix}. \quad (15)$$

We have $\mathbf{X} = \mathbf{AS}$. We now use Algorithm 1 to estimate the mixing matrix. Note that the data set is artificial, and, hence, it is not necessary to perform Step 1. The estimated mixing matrix $\mathbf{E}$ is shown in (16) at the bottom of the page.

Comparing the columns of $\mathbf{E}$ with that of matrix $\mathbf{A}$, we can see: 1) All columns of $\mathbf{A}$ are estimated very well; and 2) there exist several redundant vectors in $\mathbf{E}$. By removing these redundant vectors, we can easily have an estimation of $\mathbf{A}$. It is demonstrated that Algorithm 1 works well here, although only 8% of the $\mathbf{S}$ column vectors are sparse (of which only one entry is nonzero). In this case, it is impossible to use the K-means clustering algorithm for satisfactory estimation.

The following matrix is an estimation of mixing matrix $\mathbf{A}$ obtained by the K-means clustering algorithm

$$\begin{bmatrix} -0.5741 & 0.8917 & 0.3700 & -0.2303 \\ 0.0020 & 0.0882 & -0.0297 & 0.9686 \\ 0.8188 & 0.4440 & -0.9286 & -0.0940 \end{bmatrix}. \quad (17)$$

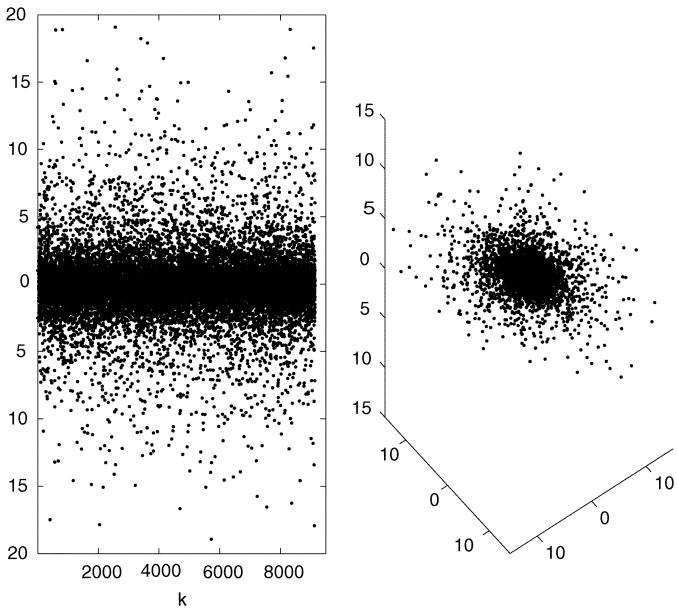

Fig. 6. (Left) Plot of all entries of matrix $(\tilde{\hat{\mathbf{X}}})$ obtained in step 3 of Algorithm 1. (Right) 3-D scatter plot of data matrix $\hat{\mathbf{X}}$, in which one point corresponds to a column value of the matrix. (Example in the Appendix).

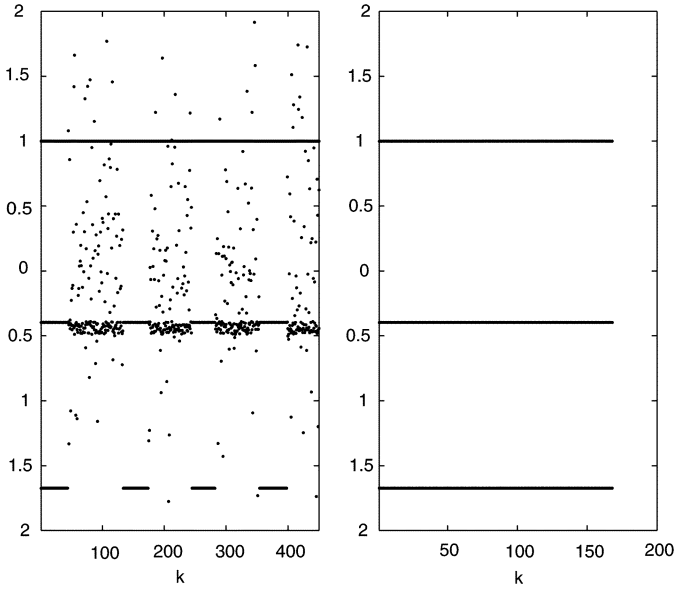

Fig. 7. (Left) Plot of all entries of matrix $(\bar{\tilde{\mathbf{X}}}_{j_k})$ obtained in step 3.2.2 of Algorithm 1. (Right) Plot of all entries of an reduced matrix; $(\bar{\tilde{\mathbf{X}}}_p^{(j_k)})$ obtained in step a2 of algorithm 1. (Example in the Appendix).

Comparing the estimation in (17) and $\mathbf{A}$, we find that the K-means clustering algorithm fails to estimate the mixing matrix.

The left subplot of Fig. 6 shows all entries of matrix $(\tilde{\hat{\mathbf{X}}})$ produced in Step 3 of Algorithm 1, and the right subplot shows

$$\mathbf{E} = \begin{bmatrix} 0.7238 & 0.8412 & 0.7240 & -0.0295 & -0.1728 & -0.0299 & -0.1728 & -0.0294 \\ 0.4087 & -0.5036 & 0.4088 & 0.8304 & 0.3628 & 0.8304 & 0.3628 & 0.8305 \\ -0.5560 & 0.1969 & -0.5555 & 0.5563 & 0.9157 & 0.5563 & 0.9157 & 0.5563 \end{bmatrix}. \quad (16)$$

the three–dimensional (3-D) scatter plot of its column values. Cursory inspection of these two plots reveals that no cluster is visible. Especially in the right subplot, the directions of the five mixing matrix columns are hidden due to the fact that only 2% of the total points represent each direction.

The left subplot of Fig. 7 shows all entries of a matrix $(\tilde{\tilde{\mathbf{X}}}_{j_k})$ obtained in Step 3.2.2 of Algorithm 1, in which we can see that several line segments are interspersed with noise. The heights of these line segments correspond to the entries of the first column of mixing matrix $\mathbf{A}$ up to a scale. Since we need a more precise estimation, we continue the procedure to further reduce matrix $\tilde{\tilde{\mathbf{X}}}_{j_k}$.

The right subplot of Fig. 7 shows all entries of a matrix $(\tilde{\tilde{\mathbf{X}}}_p^{(j_k)})$ obtained in Step a2 of Algorithm 1, in which we can see three "clear" line segments. These line segments are formed by three rows of matrix $\tilde{\tilde{\mathbf{X}}}_p^{(j_k)}$ with their heights corresponding to the three entries of the first column of mixing matrix $\mathbf{A}$ up to a scale. The normalized mean column vector of matrix $\tilde{\tilde{\mathbf{X}}}_p^{(j_k)}$ is an estimated column of the mixing matrix $\mathbf{A}$ (one of the columns of $\mathbf{E}$).

ACKNOWLEDGMENT

The authors would like to thank Dr. S. Shishkin for providing us with experimental data and useful comments.

REFERENCES

[1] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[2] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 600–616, Mar. 1997.

[3] M. S. Lewicki and T. J. Sejnowski, "Learning overcomplete representations," *Neural Comput.*, vol. 12, no. 2, pp. 337–365, 2000.

[4] B. A. Olshausen, P. Sallee, and M. S. Lewicki, "Learning sparse image codes using a wavelet pyramid architecture," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2001, vol. 13, pp. 887–893.

[5] D. L. Donoho and M. Elad, "Maximal sparsity representation via l^1 minimization," in *Proc. Nat. Acad. Sci.*, vol. 100, 2003, pp. 2197–2202.

[6] M. Zibulevsky and B. A. Pearlmutter, "Blind source separation by sparse decomposition," *Neural Comput.*, vol. 13, no. 4, pp. 863–882, 2001.

[7] S. A. Cruces-Alvarez, A. Cichocki, and S. Amari, "From blind signal extraction to blind instantaneous signal separation: Criteria, algorithms, and stability," *IEEE Trans. Neural Netw.*, vol. 15, no. 4, pp. 859–873, Jul. 2004.

[8] L. Zhang, A. Cichocki, and S. Amari, "Self-adaptive blind source separation based on activation functions adaptation," *IEEE Trans. Neural Netw.*, vol. 15, no. 2, pp. 233–244, Mar. 2004.

[9] W. Lu and J. C. Rajapakse, "Approach and applications of constrained ICA," *IEEE Trans. Neural Netw.*, vol. 16, no. 1, pp. 203–212, Jan. 2005.

[10] T. W. Lee, M. S. Lewicki, M. Girolami, and T. J. Sejnowski, "Blind source separation of more sources than mixtures using overcomplete representations," *IEEE Signal Process. Lett.*, vol. 6, no. 4, pp. 87–90, Apr. 1999.

[11] M. Girolami, "A variational method for learning sparse and overcomplete representations," *Neural Comput.*, vol. 13, no. 11, pp. 2517–1532, 2001.

[12] Y. Q. Li, A. Cichocki, and S. Amari, "Sparse representation and blind source separation," *Neural Comput.*, vol. 16, no. 6, pp. 1193–1234, 2004.

[13] ——, "Sparse component analysis for blind source separation with less sensors than sources," in *Proc 4th Int. Symp. Independent Component Analysis Blind Source Separation (ICA BSS)*, 2003, pp. 89–94.

[14] S. Makeig, M. Westerfield, T. P. Jung, S. Enghoff, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Dynamic brain sources of visual evoked responses," *Science*, vol. 295, pp. 690–694, 2002.

[15] S. Makeig, T. P. Jung, A. J. Bell, D. Ghahremani, and T. J. Sejnowski, "Blind separation of auditory event-related brain responses into independent components," in *Proc. Nat. Acad. Sci.*, vol. 94, 1997, pp. 10 979–10 984.

[16] T. P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects," *Clinical Neurophysiol.*, vol. 111, pp. 1745–1758, 2000.

[17] S. Tong, A. Bezerianos, J. Paul, Y. Zhu, and N. Thakor, "Removal of ECG interference from the EEG recordings in small animals using independent component analysis," *J. Neurosci. Methods*, vol. 108, pp. 11–17, 2001.

[18] L. Parra, C. Alvino, A. Tang, B. Pearlmutter, N. Yeung, A. Osman, and P. Sajda, "Single-trial detection in EEG and MEG: Keeping it linear," *Neurocomput.*, vol. 52–54, pp. 177–183, Jun. 2003.

[19] L. Parra and P. Sajda, "Converging evidence of linear independent components in EEG," in *Proc. 1st Int. IEEE EMBS Conf. Neural Eng.*, 2003, pp. 525–528.

[20] E. Rodriguez, N. George, J. P. Lachaux, J. Martinerie, B. Renault, and F. J. Varela, "Perception's shadow: Long-distance synchronization of human brain activity," *Nature*, vol. 397, pp. 430–433, 1999.

[21] F. Varela, J. P. Lachaux, E. Rodriguez, and J. Martinerie, "The brainweb: Phase synchronization and large-scale integration," *Nature Rev., Neurosci.*, vol. 2, pp. 229–239, 2001.

[22] W. Krause, H. Gibbons, and B. Schack, "Concept activation and coordination of activation procedure require two different networks," *NeuroReport 9*, pp. 1649–1653, 1998.

[23] S. Weiss and P. Rappelsberger, "Long-range EEG synchronization during word encoding correlates with sucessful memory performance," *Congn. Brain Res.*, vol. 9, pp. 299–312.

[24] C. Allefeld and J. Kurths, "Multivariate phase synchronization analysis of EEG data," *IEICE Trans. Fundam.*, vol. E86 A, no. 9, 2003.

[25] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis," *Brain Res. Rev.*, vol. 29, pp. 169–195, 1999.

[26] R. Q. Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger, "Performance of different synchronization measures in real data: A case study on EEG signals," *Phys. Rev. E*, vol. 65, 2002.

[27] S. R. Campbell, D. Wang, and C. Jayaprakash, "Synchronization rates in classes of relaxation oscillators Campbell," *IEEE Trans. Neural Netw.*, vol. 15, no. 5, pp. 1027–1038, Sep. 2004.

[28] J. P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela, "Measuring phase synchrony in brain signals," *Hum. Brain Mapping*, vol. 8, pp. 194–208, 1999.

[29] M. Le Van Quyen, J. Foucher, J. P. Lachaux, E. Rodriguez, A. Lutz, J. Martinerie, and F. J. Varela, "Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony," *J. Neurosci. Methods*, vol. 111, pp. 83–98, 2001.

[30] B. Schack, N. Vath, H. Petsche, H. G. Geissler, and E. Moller, "Phase-coupling of theta-gamma EEG rhythms during short-term memory processing," *Int. J. Psychophysiol.*, vol. 44, pp. 143–163, 2002.

[31] M. J. Hogan, G. R. J. Swanwick, J. Kaiser, M. Rowan, and B. Lawlor, "Memory-related EEG power and coherence reductions in mild Alzheimer's disease," *Int. J. Psychophysiol.*, vol. 49, pp. 147–163, 2003.

[32] P. M. Lahteenmaki, C. M. Krause, L. Sillanmaki, T. T. Salmi, and A. H. Lang, "Event-related alpha synchronization/desynchronization in a memory-search task in adolescent survivors of childhood cancer," *Clinical Neurophysiol.*, vol. 110, pp. 2064–2073, 1999.

[33] G. Pfurtscheller and W. klimesch, "Functional topography during a visuoverbal judgment task studied with event-related desynchronization mapping," *J. Clinical Neurophysiol.*, vol. 9, pp. 120–131, 1992.

[34] H. Supér, C. van der Togt, H. Spekreijse, and V. A. F. Lamme, "Internal state of minkey primary visual cortex (V1) predicts figure-ground perception," *J. Neurosci.*, vol. 23, no. 8, pp. 3407–3414, 2003.

[35] P. J. Durka, "Adaptive time-frequency parameterization of epileptic spikes," *Phys. Rev. E*, vol. 69, pp. 051 914–051 914, 2004.

[36] P. J. Durka and K. J. Blinowska, "A unified time-frequency parametrization of EEG," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 5, pp. 47–53, Sep./Oct. 2001.

[37] Y. Li, S. Amari, A. Cichocki, D. W. C. Ho, and S. Xie, "Underdetermined blind source separation based on sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 423–437, Feb. 2006.

Yuanqing Li was born in Hunan Province, China, in 1966. He received the B.S. degree in applied mathematics from Wuhan University, Wuhan, China, in 1988, the M.S. degree in applied mathematics from South China Normal University, Guangzhou, China, in 1994, and the Ph.D. degree in control theory and applications from South China University of Technology, Guangzhou, China, in 1997.

Since 1997, he has been with South China University of Technology, where he became a Full Professor in 2004. He spent a few years at the Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Saitama, Japan, as a Researcher. He is the author or coauthor of more than 50 scientific papers in journals and conference proceedings. His research interests include automatic control, blind signal processing, neural networks, brain-computer interface, neural coding, and neural decoding.

Andrzej Cichocki was born in Poland. He received the M.Sc. (with honors), Ph.D., and Habilitate Doctorate (Dr.Sc.) degrees, all in electrical engineering, from Warsaw University of Technology, Warsaw, Poland, in 1972, 1975, and 1982, respectively.

Since 1972, he has been with the Institute of Theory of Electrical Engineering and Electrical Measurements at the Warsaw University of Technology, where he became a Full Professor in 1991. He is the coauthor of three books: *Adaptive Blind Signal and Image Processing* (Wiley: New York, 2003 with Professor Shun-ichi Amari), *MOS Switched-Capacitor and Continuous-Time Integrated Circuits and Systems* (Springer-Verlag: New York, 1989, with Professor Rolf Unbehauen) and *Neural Networks for Optimization and Signal Processing* (Wiley and Teubner Verlag, 1993/94, with Professor Rolf Unbehauen) and coauthor of more than 150 papers. Two of his books have been translated to Chinese and other languages. He spent a few years at the University Erlangen-Nuernberg, Germany, as Alexander Humboldt Research Fellow and Guest Professor working in the area very large scale integration (VLSI) of electronic circuits, artificial neural networks, and optimization. He conducted and realized several successful research projects. In 1996–1999, he worked as a Team Leader of the Laboratory for Artificial Brain Systems at the Frontier Research Program RIKEN, Japan, in the Brain Information Processing Group. Since 1999, he has been the Head of the Laboratory for Advanced Brain Signal Processing in the Brain Science Institute RIKEN, Japan. Since January 1998, he has been the Associate Editor of IEEE TRANSACTION ON NEURAL NETWORKS. His current research interests include biomedical signal and image processing (especially blind signal/image processing), neural networks and their applications, learning theory and robust algorithms, generalization and extensions of independent and principal component analysis, optimization problems and nonlinear circuits, and systems theory and their applications.

Dr. Cichocki is a member of several international scientific committees.

Shun-Ichi Amari (M'71–SM'92–F'94) was born in Tokyo, Japan, on January 3, 1936. He received the Dr. Eng. degree in mathematical engineering from the Graduate School of the University of Tokyo, Tokyo, Japan, in 1963.

He worked as an Associate Professor at Kyushu University and the University of Tokyo, and then a Full Professor at the University of Tokyo, and he is now Professor-Emeritus. He serves now as Director of RIKEN Brain Science Institute, Saitama, Japan. He has been engaged in research in wide areas of mathematical engineering, such as topological network theory, differential geometry of continuum mechanics, pattern recognition, and information sciences. In particular, he has devoted himself to mathematical foundations of neural network theory, including statistical neurodynamics, dynamical theory of neural fields, associative memory, self-organization, and general learning theory. Another main subject of his research is information geometry initiated by himself, which applies modern differential geometry to statistical inference, information theory, control theory, stochastic reasoning, and neural networks, providing a new powerful method of information sciences and probability theory.

Dr. Amari is the President of Institute of Electronics, Information and Communication Engineers, Japan, and Past President of International Neural Networks Society. He received the Emanuel A. Piore Award and the Neural Networks Pioneer Award from the IEEE, the Japan Academy Award, the C & C Award, among many others.