ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 34 (2007) 1519–1534

Localization of individual area neuronal activity

N. Hironaga* and A.A. Ioannides

Laboratory for Human Brain Dynamics, RIKEN Brain Science Institute (BSI), Wako-shi, Saitama 351-0198, Japan
Kyushu Institute of Technology, Department of Brain Science and Engineering, Wakamatsu Ku, 2-4 Hibikino, Kitakyushu, 808-0196, Japan

Received 21 November 2005; revised 26 September 2006; accepted 17 October 2006
Available online 21 December 2006

A family of methods, collectively known as independent component analysis (ICA), has recently been added to the array of methods designed to decompose a multi-channel signal into components. ICA methods have been applied to raw magnetoencephalography (MEG) and electroencephalography (EEG) signals to remove artifacts, especially when sources such as power line or cardiac activity generate strong components that dominate the signal. More recently, successful ICA extraction of stimulus-evoked responses has been reported from single-trial raw MEG and EEG signals. The extraction of weak components has often been erratic, depending on which ICA method is employed and even on what parameters are used. In this work, we show that if the emphasis is placed on individual "independent components," as is usually the case with standard ICA applications, differences in the results obtained for different components are exaggerated. We propose instead the reconstruction of regional brain activations by combining tomographic estimates of individual independent components that have been selected by appropriate spatial and temporal criteria. Such localization of individual area neuronal activity (LIANA) allows reliable semi-automatic extraction of single-trial regional activations from raw MEG data. We demonstrate the new method with three different ICA algorithms applied to both computer-generated signals and real data. We show that LIANA provides almost identical results with each ICA method despite the fact that each method yields different individual components.

Keywords: MEG; ICA; Regional activation; Single-trial analysis

Introduction

Raw magnetoencephalography (MEG) and electroencephalography (EEG) signals are contaminated by interference from ambient noise and unrelated signals generated by the subject. Although we are not interested in these noise components, it is important to understand their nature so that they can more effectively be identified and eliminated. Ambient noise includes interference from commercial power lines (operating at 50 Hz in our laboratory), its harmonics, and DC drift, as well as electronic noise from thermal sensor instabilities and the acquisition system. Other unwanted components include strong signals from the heart, facial muscles, eye activity, and so on. In most cases, we are interested only in part of the brain activity—very often a small fraction of what goes on. If a convenient time reference can be defined—for example the onset of repeated stimulation—then averaging can be used to capture brain activity that is time-locked to the stimulus onset.

Since the mid-1990s, independent component analysis (ICA) has been applied in MEG/EEG studies too numerous to adequately summarize here. In brief, early applications focused on the identification and elimination of artifacts (Vigario, 1997; Barros et al., 1998; Jung et al., 2000a,b; Joyce et al., 2004; Barbati et al., 2004; Wallstrom et al., 2004). Automatic identification of artifacts was suggested to be the main advantage of ICA-based methods compared to the more widely used principal component analysis (PCA). The application of ICA to event-related studies was initially introduced by the Salk Institute group (Makeig et al., 1997) and soon after by Vigario and colleagues (1998, 1999) in Finland using FastICA. Source localization of ICA components for medical application was initially proposed for epileptiform discharges (Kobayashi et al., 1999, 2002a,b; Ossadtchi et al., 2004). The Salk Institute group has reported ICA applications at both the average (Makeig et al., 1997, 1999) and single-trial level (Jung et al., 2001; Makeig et al., 2002; Delorme et al., 2002) using INFOMAX. Applying ICA to averaged data has also been discussed by the same group (Delorme and Makeig, 2004). Tang and colleagues (2002a,b, 2005) reported an improvement in response onset detection in single trials using second-order blind identification (SOBI). Lee and colleagues (2003) focused on the relationship between responses evoked by stimuli and that of the ongoing background activity using FastICA.

MEG and EEG are the only techniques offering non-invasive measures of macroscopic brain activity with millisecond temporal resolution. In the case of MEG, strong claims have also been made for high spatial resolution (Moradi et al., 2003), but the prevailing view is that the spatial resolution of MEG is not as good as that of

* Corresponding author. Laboratory for Human Brain Dynamics, RIKEN Brain Science Institute (BSI), 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan. Fax: +81 48 467 9731.
E-mail address: hironaga@brain.riken.go.jp (N. Hironaga).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2006.10.030

functional magnetic resonance imaging (fMRI). A number of methods have been proposed for identifying the generators of the MEG and EEG signal. They can be divided into methods assuming point-like generators (Scherg, 1990) or distributed sources. The majority of distributed-source solutions rely on linear inverse algorithms and they are summarized in (Sarvas, 1987; Hamalainen et al., 1993). Here we use a non-linear distributed-source reconstruction method (Ioannides et al., 1990). By time locking the measurements to a stimulus, repeating them many times, and averaging the responses, one usually obtains a signal that is characterized by early peaks that often contain dominant contributions from at least one, and sometimes a few, focal generators. Simple models can then localize these generators accurately. Research conducted with non-invasive electrophysiology has been dominated by averaging signals with trials aligned to some external (e.g., a stimulus) or internal (e.g., saccadic onset) event. The resulting average MEG or EEG signals are known as event-related potentials (ERPs) or fields (ERFs), respectively. We will refer to both collectively as ERF(P)s hereafter. Recent studies have demonstrated that ERF(P) peak responses reflect, at least in part, event-related reorganization and synchronization of ongoing oscillations (Liu et al., 1998; Makeig et al., 2002; Laskaris et al., 2003; Yeung et al., 2004). The background oscillations are often generated in polymodal areas (Laskaris et al., 2003), well-separated from primary sensory areas where responses can be adequately described by focal generators. Averaging makes limited use of MEG's time sensitivity, reducing the huge dynamic interchanges between brain areas that are evident in the raw signal to small and possibly insignificant stereotyped responses (Laskaris et al., 2003).

In our earlier studies, we faced the problem by applying an optimal, non-linear tomographic analysis to the MEG signal of each time slice of each single trial. Regional activations were then extracted from the solutions and treated as a pattern in space and time. This computationally intensive procedure is realized in four steps: First, the time evolution of the signal is used to identify and eliminate sensors and trials with obvious large artifacts. Second, methods like PCA or ICA are used to identify and remove any remaining contributions to the signal from artifacts with a fairly stereotyped pattern (e.g., eye movements and heart and muscle activity). In the third step, the "clean" spatial patterns of signal topography for each time slice of each single trial is used to tomographically reconstruct the best estimate of activity throughout the brain. Finally, statistical analysis is performed on the large volume of single-trial solutions. This approach has produced new insights (Ioannides, 2001; Ioannides and Fenwick, 2005; Ioannides et al., 2004, 2005; Laskaris et al., 2003), but at a huge computational cost. We introduce here an alternative three-step method. The new approach combines optimal tomographic analysis and ICA to extract results similar to a full tomographic analysis of single-trial data, but at a much lower computational cost. Previously, a tomographic analysis was performed on each time slice of data. Instead, we propose to reconstruct the source distribution by combining the tomographic maps of a subset of components using weights determined by the activation time courses of each IC in the subset. In the limit of only one IC, the method reduces to the standard usage of ICA. In the first step of the new approach, we use ICA to identify large artifacts and useful components. We then focus on the localization of the useful components and use the resulting tomographic descriptions of IC activity to select the components that contribute significantly to the activity of a given region of interest (ROI). The time courses corresponding to the selected ICs, as calculated by ICA, are then used to compute the contribution for the grid points within a specified ROI and along the spatial direction given by the selected ICs. ICA could be replaced by any other method of decomposing the components of a mixed signal, but in this context ICA is preferred because the blind separation of components allows the signal to determine what the dominant contributions are. The individual elements that are used in the new method are not new, since many authors have considered different ways of extracting ICA components and have proposed the localization of individual components. What is novel is the use of a combination of ICA components selected for their contribution to a specific ROI so that regional activations can be mapped in the subspace defined by the selected components.

It has repeatedly been shown that ICA can detect power line artifacts, sensor noise, evoked responses, oscillatory activations, eye movements, blinks, cardiac artifacts, and muscle artifacts. A closer look at what is captured by an IC shows a rather capricious capability, especially when many components are present with fairly similar strengths. Each component corresponds to a mathematical partitioning of the signal into a set of magnetic field generators (Tang et al., 2002a). There is no theoretical requirement that any one of the mathematical components extracted by ICA correspond to a single underlying generator, such as a focal activation. In principle, one underlying generator may be recovered by several ICs (Jung et al., 2001; Lee et al., 2003). In this paper, we first investigate the relationship between ICs extracted by different popular ICA methods, and their underlying generators. Several recent ICA studies have demonstrated components that appear to capture activity from a single underlying generator which is relatively free of contamination (Makeig et al., 2002; Tang et al., 2005; Debener et al., 2005). In almost every case, a strong component with narrow-band filter is singled out for the analysis. However, a relatively weak source can spread into multiple components, and individual ICs may contain contributions from different sources. We thus introduce the new three-step method for localizing individual area neuronal activity (LIANA). LIANA recovers regional brain activations from the raw signal and the ICs that have been identified as significant contributors to the ROI's activity.

In the next section, we provide an overview of ICA and describe the data we have used for the tests, but leave to the Appendix the description of the individual algorithms that we have included in this work. The detailed steps of LIANA are introduced next, followed by some comments on localization of ICs. In the following section, we test LIANA using three distinct ICA algorithms and three separate types of data sets. The first set of data uses MEG signals from resting (control) condition. The second set adds computer-generated data with known position and well-defined generator activity to the resting data signal. Finally, data from an actual somatosensory experiment are analyzed with LIANA to calculate the generator activity both for the averaged responses and the single-trial data.

Materials and methods

Data acquisition

The data were acquired using the VSM whole-head 151-channel system (VSM MedTech Ltd., Vancouver, Canada) as part of a series of somatosensory experiments performed in our

laboratory and approved by the RIKEN Ethical Committee. Six healthy right-handed subjects participated in these experiments. The protocol was explained to each subject who then signed an informed consent form. The data used in this paper are from two male subjects who at the time of the experiment were 28 and 29 years old. These two subjects also participated in additional runs with high sampling rate data collections. One of these subjects was of average height and the other taller than average. We performed 1 min of resting data collection before the somatosensory experiment. During resting data collection, subjects were required to relax but fixate their gaze as carefully as possible. The somatosensory stimuli were delivered using a GRASS stimulator (GRASS S8800, Astro-Med, Inc., USA). Square-shaped current pulses of 0.2 ms duration were delivered to the left and right median nerves at the wrist. The strength of stimulation was well above sensory threshold but just below motor threshold, and the interstimulus interval varied randomly within a 200 ms window centered on 1 s. In each of four active runs lasting about 4 min each, the left and right wrists were stimulated in a pseudorandom order. In each run, the total number of trials for each wrist was 120. The results were similar across subjects. Since the ICA and LIANA methodologies are the primary interest here, we will show the results from just one subject's resting data and another subject's real data analysis. In addition to the stimulation electrodes, we attached auxiliary channels for monitoring eye movement and heart activity. Separate electrode pairs placed above, below, and to the left and right of the eyes monitored the horizontal and vertical electrooculogram (EOG). To monitor the subject's heart function, we used 4 pairs of electrocardiogram (ECG) electrodes placed on the left and right wrists, left and right ankles, and lead V2.

We used the highest sampling frequency (2083 Hz) allowed by our system for simultaneous recording of all MEG and auxiliary channels. A 600 Hz low-pass (LP) firmware filter was applied during recording to avoid aliasing problems. Before MEG acquisition, 3D camera measurement (VIVID 700, Minolta Co. Ltd., Japan) and 3D digitization using the Polhemus system (FastTrack, Polhemus, Colchester, USA) were performed for accurate co-registration (Hironaga et al., 2002). DC removal, gradient formation, and high-pass (HP) 2.084 Hz filtering were applied to all data sets after recording.

ICA basics

We consider two general types of blind source separation (BSS) algorithms, which for the purposes of this paper we treat as two classes of ICA. The first set of methods is based on probability density and high-order statistics, negentropy (Comon, 1994), INFOMAX with natural gradient (Bell and Sejnowski, 1995; Amari, 1998), INFOMAX-EXT (Lee et al., 1999), FastICA (Hyvarinen and Oja, 1997; Hyvarinen, 1999), and joint approximate diagonalization of eigenmatrices (JADE) (Cardoso and Souloumiac, 1996). The second set uses time-delayed correlation-based methods (Molgedey and Schuster, 1994). We emphasize a specific implementation of the second set of methods, SOBI (Belouchrani et al., 1997), which has recently become popular thanks to a number of successful applications (Tang et al., 2002a,b, 2005; Kishida et al., 2003; Sander et al., 2005). In the end, we employed INFOMAX, FastICA, and SOBI to keep down the computational cost.

In general, ICA decomposition starts from the following model,

$$x = As \tag{1}$$

where $x=\{x_1(t),x_2(t),\ldots,x_n(t)\}$, a matrix of recorded sensor signals with n the number of MEG channels, and t ranging from 1 to m, the number of sampling points. A is a square matrix of dimension n, and $s=\{s_1(t),s_2(t),\ldots,s_n(t)\}$, the matrix of independent source signals. The solution is sought as a linear transformation of the detected signals,

$$y = Wx \tag{2}$$

where W is a matrix to be determined, generally called a separating matrix. Without any knowledge or constraint on the source configuration, the task is to reconstruct y (which is assumed to be equivalent to s). The core part of ICA is to find a square matrix W that linearly inverts the mixing process. Ideally, W must be A^{-1}; however, because of scale and permutation problems, it cannot be uniquely defined. The algorithm leads to W which satisfies,

$$WA = PD \tag{3}$$

where P is a permutation matrix, D is a diagonal matrix. W is estimated by following an iteration process according to the definition of independence. The ICA problem is simplified if the input signal is uncorrelated and the contribution from different channels is normalized to avoid overemphasizing the contribution from the strongest channels. It is therefore common to apply whitening as a preprocessing step, for example, by transforming the input to,

$$z = Vx \tag{4}$$

$$E\{zz^T\} = I \tag{5}$$

Here and in what follows, I denotes a unit matrix, and E the expectation operator. In Eq. (4), V denotes a transformation matrix, and z in Eq. (5) is the transformed matrix (vector), which can be calculated simply by eigenvalue decomposition as follows:

$$V = \Lambda^{-1/2}U^T \tag{6}$$

where U is the matrix whose columns are the unit-norm eigenvectors of the covariance matrix of x, and Λ is the diagonal matrix of the eigenvalues of the covariance matrix. The ICA algorithms employed in this paper and implementation parameters will be described in the Appendix.

Localization of independent components

The final output of each ICA method is an estimate of the separating matrix, W, and its approximate inverse which represents an estimate of the mixing matrix A. Given such an estimate for A, a set of "independent components" or ICs can be obtained, with the ith IC given by

$$[As]_i = \begin{pmatrix} a_{1i}s_i(t_1) & \cdots & a_{1i}s_i(t_m) \\ \vdots & & \vdots \\ a_{ni}s_i(t_1) & \cdots & a_{ni}s_i(t_m) \end{pmatrix} = (a_{*_i})(s_i(t)) \tag{7}$$

where $t=(t_1,,,t_m)$, m is the number of sample points, and n is the number of MEG channels. It is obvious that the sensor loading for each component is time independent and is specified for this component as a_{*i}, for the ith column of matrix A. Given the sensor loading, many choices can be used for estimating the corresponding generator distribution. We avoid the usual choice of equivalent current dipoles (ECD) because it restricts the solution to one location. We want to allow for the possibility that a given component may describe distributed activity and/or focal activations in more than one location and to do so in a continuous way so that the contributions from different ICs can be combined according to their strength at each point in space and time. Another advantage of most distributed source solutions, including the MFT method which we will describe next, is that they allow the use of a regularization parameter that balances the compromise between accurate reconstruction and over-fitting of the data (i.e., limiting the influence of noise on the data).

Magnetic field tomography (MFT)

In principle, any distributed source solution could be used to localize the individual ICs. We use magnetic field tomography (MFT) (Ioannides et al., 1990) because this method is optimal for estimating activity in the brain tomographically from each single time slice of MEG data (Taylor et al., 1999), or from the sensor loading of each IC. This claim is based on close examination of the mathematical properties of the lead fields, which has led to a generalized version of MFT that can be related to most other popular distributed source methods (Taylor et al., 1999; Ioannides and Taylor, 1999). The key point is rather subtle. Although the "forward" relationship from the generators to the data is linear, a mathematically consistent formulation of the inverse relationship expressing the generators in terms of the data is non-linear (Taylor et al., 1999). It is the use of the optimal non-linearity that sets MFT apart from other methods (Taylor et al., 1999; Ioannides and Taylor, 1999). Solutions produced by linear methods are too smooth and they are biased towards the sensors. This bias towards the sensors can be corrected by expressing the unknown current density vector as a linear sum of the sensitivity profile of the sensors, modulated by a depth weighting factor, $w(\mathbf{r},J)$, where $\mathbf{r}$ is the vector defining the current location in space, and J is the current density modulus at that point in space. Dependence of w on J leads to a non-linear system of equations that is in general difficult to solve. A linear framework requires that w be independent of J, i.e., that $w=w(\mathbf{r})$. Although such an "ansatz" seems reasonable and is widely used, it has no justification other than the fact that it leads to a linear system of equations and hence to computational simplicity. The underlying physics provides no foundation for a J-independent w. Physics allows only the local direction (unit vector) of the current density to be expressed in terms of a J-independent weight factor, i.e., $w=|J(\mathbf{r})|*w_0(\mathbf{r})$. This is exactly what standard MFT assumes, with the form of this *a priori* weight, $w_0(\mathbf{r})$, determined by training with computer-generated data (Taylor et al., 1999). The mathematical structure of the expansion of the current density in terms of lead fields in standard MFT allows for sharp transitions in intensity with no mathematical singularities because the expansion of the primary current density depends only on the direction. One or more sharp transitions (as would be required by point sources) are obtained just as easily as smooth distributions because no huge expansion coefficients are involved. The details in the data determine whether point-like generator(s) or smooth distributions are recovered, and the data alone determine the strength of activation at each point in space. In this sense, MFT makes maximal use of the information in the data with minimal *a priori* assumptions. The non-linearity of the MFT algorithm requires that the method be applied independently for each time slice of data, thus making huge computational demands. The advantage, on the other hand, is that it leaves the question of statistical analysis open so that any statistical analysis can be performed at the level of the estimates of activity in the brain coordinates. For exactly the same reasons, MFT is optimal for combining distributed source solutions within a subset of the space defined by ICA analysis; the tomographic description of each IC is derived independently and under minimal assumptions.

From a purely pragmatic point of view, applications of MFT to real MEG data have consistently produced solutions with surprisingly good accuracy. At the cortical level, an accuracy of a few millimeters has been demonstrated (Moradi et al., 2003). Recently, good localization was also demonstrated for deeper structures including the amygdala and brainstem (Ioannides et al., 2004) and cerebellum (Ioannides and Fenwick, 2005; Ioannides et al., 2005).

Signal recovery and generators

The immediate task of many signal processing methods including PCA and ICA analysis is to provide a decomposition of the signal into components that relate to the underlying generators better than the original time courses of individual channels. The overall goal is an accurate estimation of the number, distribution in space, and time course of different generators, and the way their activity is modulated by stimuli and task demands. Fig. 1 provides a schematic view of what is involved. We exclude from further analysis the occasional data segment that is contaminated by large artifacts, as well as the signal from noisy or flat sensors. The remaining signal is then analyzed with ICA. Then, visual inspection, threshold level, or correlation with some auxiliary channel (e.g., ECG) is used to identify ICA components that are dominated by unwanted interference. The remaining components are used to reconstruct the signal and different algorithms can be used to localize its generators.

In any classification scheme, the aim is to analyze the structure of brain activity in space and time. Just as there is no justification in assuming that each peak in the average signal corresponds to the activity of one generator, there is no *a priori* justification either for a one-to-one correspondence between ICs and generators. Furthermore, in the case of ICA, the linear unmixing procedure requires that the number of ICs be equal to the number of input MEG channels. Of course, there is no reason for this exact number of generators to be activated during any given finite interval. The fact that different components are produced by different ICA algorithms indicates that the mapping between ICs and real generators will not be one-to-one. A one-to-one mapping is possible if there is at most the same number of generators as independent sensors, and the specific assumptions made by any one ICA method used are satisfied by all generators in the brain. Whatever ICA method is used, it is likely to be appropriate for only a fraction of the generators. Since the contribution from all generators is mixed in the signal, the effectiveness of the unmixing ICA process will be compromised even for generators that have the appropriate time course structure.

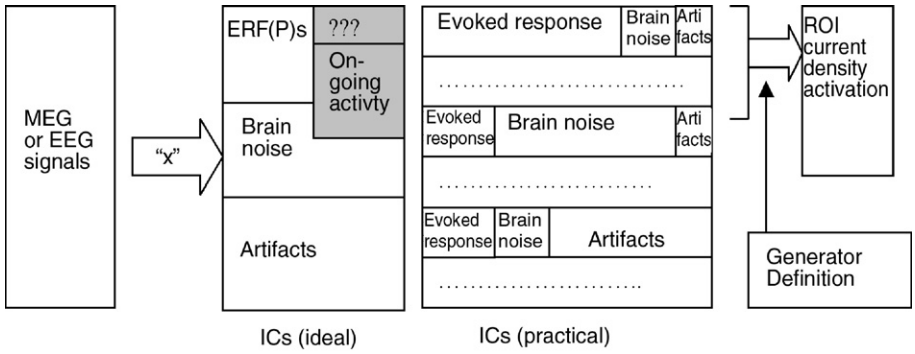

Fig. 1. Component and generator definition, with "*x*" denoting the total set of processes applied to the ERF(P)s signals leading to a new decomposition of the signal. Examples of pre-processing operations include removal of noisy sensors and segments with contamination by huge artifacts, filtering and whitening. Examples of operations aiming specifically at producing new component classification include selection of special sets of trials. The ideal output of process "*x*" is conceptually separated into three types. In the first type, ideal ERF(P) contains the signal of interest as it is generated by the brain, possibly with some signal from other generators that could not be separated (the "???" symbols acknowledge this uncertainty). We specifically mark contributions from ongoing background rhythms because even if they are largely eliminated by some of the processes, their indirect effect may still be present. The second type of ideal set is composed of components that may be generated by the brain but are of no interest to the aims of the current study and hence are labeled "brain noise." Finally, the last type contains artifacts from the body or environment. In practice, such an output of process "*x*" is a set of components, each containing signal of interest, brain noise, and contribution from artifacts. Some *a priori* knowledge of what the signal of interest should be is often derived from assumptions about the generator properties, and it is used to restrict the actual components to be used for reconstructing a point-generator description or a continuous current density estimate of the activity.

LIANA

Motivated by these considerations, we propose a method that requires only that whatever ICA method is used, it should produce an approximate separation of components, partly based on location, and partly based on time structure. The new method, LIANA, starts from an ROI specification. The ROI could be a single compact brain area (e.g., SI), a set of such areas, or a large continuous cortical region. For simplicity, we will use a single compact region and its activation time course. The specification of the ROI can be based on either anatomical (e.g., Talairach coordinates or visual identification of anatomical landmarks from MRI) or functional considerations, or a combination of the two (e.g., the strongest consistent activation close to a given sulcus). Since activity from a given brain area may be captured by multiple ICs, LIANA uses ICA only as a tool for extracting all relevant ICs with high relative strength within the pre-defined ROI. Further criteria can be imposed to allow for specific time course profiles, e.g., time locking (exact or approximate) to the onset of an external stimulus or endogenously generated output (e.g., electromyography (EMG)). The ICs identified this way are combined with appropriate weights to define the full regional brain activity of interest in each single trial. The main contrast between LIANA and previous approaches is the use of all ICs satisfying the set selection criteria (see next section) in extracting an activity profile rather than requiring ICA itself to produce ICs that correspond one-to-one with regional brain activations of interest. In almost all previous cases, target activity (thought to correspond to a real generator in the brain) was identified component-by-component by individual ICs, using mostly visual inspection of the time courses followed by averaging and localization after sensor back-projection or analysis of the frequency spectrum of individual components. LIANA makes it possible to skip detailed visual inspection. The steps making up the LIANA algorithm are shown schematically in Fig. 2.

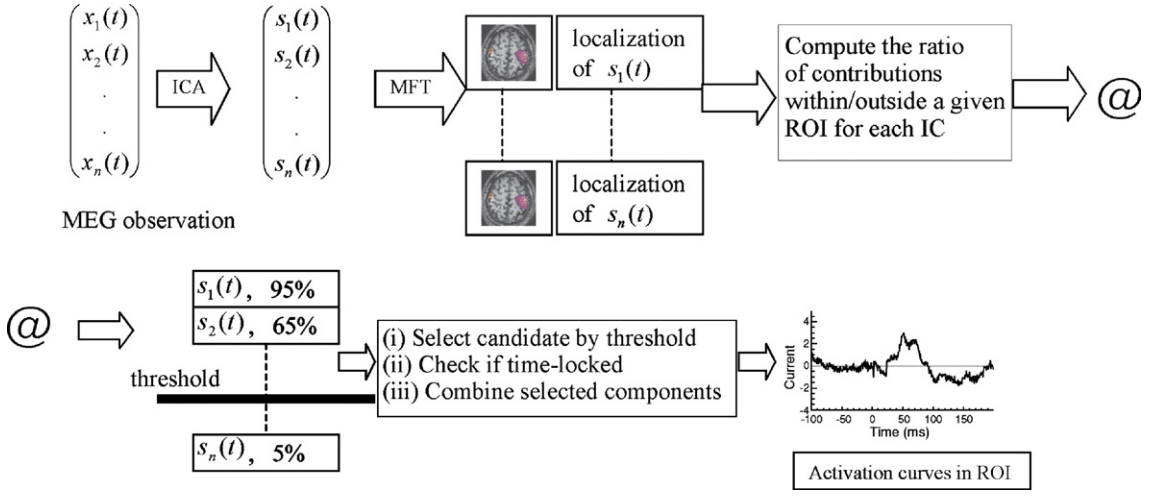

Fig. 2. Schematic flow of the steps in LIANA.

Selection of independent components in a specified ROI

A target ROI is first delineated from the subject's brain image (MRI). As a typical example, we will consider somatosensory-evoked responses and examine the activations in the SI and in the superior parietal lobule (SPL), a few centimeters posterior to SI. A preliminary identification is first made for these ROIs relative to the central sulcus. The center point and region of interest is then marked more precisely for each subject from the focal response around 20 ms (for SI), and 60 ms (for SPL), as identified by the localization of the averaged MEG signal (ICA clean data). We use a radius of 1.0 cm for each ROI. The sensor weights of each IC are then fed into the MFT algorithm to obtain a three-dimensional tomographic representation. If the *i*th IC is significantly localized within the ROI of interest, it is selected automatically using the following criteria:

$$[\underset{\substack{k\in \mathrm{ROI}\\ t\in T}}{\mathrm{Max}}\,(g_j^i(t))/\underset{\substack{k\in G\\ t\in T}}{\mathrm{Max}}(g_k^i(t))]*100\geq R, \tag{8}$$

where ROI denotes grid points within the ROI; G denotes the whole set of grid points (the entire brain); T denotes some time interval of interest (or the entire time range analyzed); and *Max* denotes the maximum-value operation over time t and grid points j, or time t and grid points k. Finally, we obtain $g_m^i(t) = |\overrightarrow{V_m^i}|s_i(t)$ with $\overrightarrow{V_m^i}$ the MFT solution (current density vector) for the ith IC at the mth grid point. The MFT solutions for each IC are also used to order ICs in decreasing order according to their Frobenius norm, so that the strongest ICs have the smallest sort number. We used both the sort number and ratio threshold of Eq. (8) with threshold levels, R, from 50% to 90%. ICs with small sort number and a large percentage of contribution to the target ROI were retained. The choice of ICs based solely on these two criteria could still be biased from contributions from a second area that is very close (e.g., the SI and motor cortices). The method allows the introduction of more criteria based, for example, on the direction of the current density or the time course of individual components. Here, in the case of the somatosensory response of interest, we identify time-locked responses by averaging each IC time course after alignment with the contralateral wrist stimulus onset.

Recombination of selected ICs

As already described, the MFT solution of the ith IC is determined from the sensor weights, a_{*i}, the ith column of (the ICA estimate of) matrix A, while the time course of the ith IC is independently described by $s_i(t)$, the ith row of matrix s. Here, we briefly describe how to reconstruct the activation curves in the ROI using multiple ICs. The (vector) activation curve of this IC at the grid point is, by definition,

$$\mathrm{ACT}_j^{\mathrm{IC}_i}(t) = \overrightarrow{V_j^i}s_i(t), \tag{9}$$

where $s_i(t)$ is the time course in Eq. (7). If more than one IC is allowed, then the activity at the grid point j at time t is simply,

$$\mathrm{ACT}_j(t) = \sum_{i\in \mathrm{ROI}} \overrightarrow{V_j^i}s_i(t), \tag{10}$$

where, ϵROI specifies that the sum runs only over ICs that have been selected for the given ROI. Finally, the activity of an ROI that includes more than one grid point and has a preferred direction along the unit vector $\hat{e}$, can be written as

$$\mathrm{ACT}_{\mathrm{ROI}}(t) = \sum_{i\in \mathrm{ROI}}\sum_j c_j(\overrightarrow{V_j^i}\cdot\hat{e})s_i(t), \tag{11}$$

where c_j is a weight factor emphasizing more contributions close to the ROI center. In this paper, we have used the following formula,

$$c_j = \frac{d_j}{\sum d_k}, d_j = \frac{1}{e^{(r_j/R)^2}}, \tag{12}$$

where r_j is the distance between the jth grid point from the center of the ROI, and R is the radius of the ROI. We note that the method easily generalizes to allow for directions varying within an ROI, e.g., following the local normal of the cortical surface.

Results

Resting data

First we analyzed the 1 min resting data collected before the somatosensory stimulation experiment to see what time courses LIANA would produce from data with just spontaneous brain activity. The most popular application of ICA to MEG/EEG is in the elimination of artifacts generated by cardiac activity and eye movements. We only examined ICs that captured cardiac activity because previous assessments of eye blink activity have been extensively discussed (Joyce et al., 2004). For average height subjects, the cardiac artifact was fairly strong and fixation was interrupted by only a few blinks with no definite eye movements. In general cardiac activity is unevenly distributed among the MEG sensors. Typically, the sensors at the base of the helmet and over temporal and occipital areas detect the strongest cardiac signal. The cardiac artifact also varies considerably from subject to subject. We consistently find that cardiac activity from tall subjects produces less contamination. In such cases, the MFT estimates are not unduly influenced by cardiac interference. A small cardiac interference is not an advantage for ICA identification. If the cardiac contribution is similar to other brain activations then it often fails to be captured by just one or few ICs, as it spreads over many ICs together with signals generated by other generators. For short subjects, the heart is closer to the sensors and a strong artifact is produced that is easily picked up by almost all ICA algorithms into one, or no more than a few ICs. Fig. 3 shows a typical example of cardiac activity from the average height subject studied in the experiment. In this case, each one of the ICA algorithms recovered one IC dominated by cardiac activity. From the inspection of the time domain (Fig. 3), the cardiac IC for FastICA (Fig. 3a) and INFOMAX (Fig. 3b) show clear detection of cardiac signal and give similar results in capturing the cardiac signal. With SOBI (Fig. 3c), there is a hint that the IC dominated by cardiac activity includes contributions from other generators. In the frequency domain (Fig. 4), a weak power line peak (50 Hz) can be seen in the SOBI result (Fig. 4c), while it is more evident in the FastICA results (Fig. 4a) and INFOMAX results (Fig. 4b). It has already been demonstrated that SOBI separates the power line artifacts into few ICs (Kishida et al., 2003). Oscillatory activity, particularly in the alpha range, is detected by all algorithms. Several components are detected with a frequency peak in the alpha range. The time domain signals of the most representative alpha IC from each method (defined as the component with the smallest sort number and distinct alpha peak) are displayed in Fig. 5a and corresponding

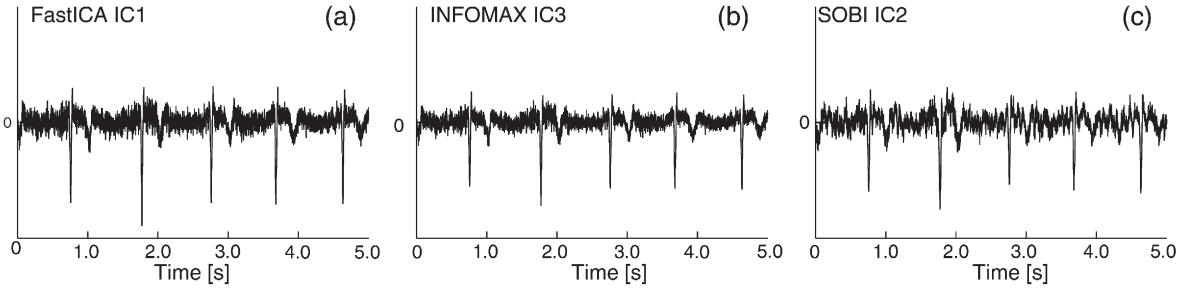

Fig. 3. The cardiac signal recovered using three ICA methods. All methods produce one IC that appears to capture well the cardiac contribution.

MFT solutions for each IC in Fig. 5b. The ICs selected for Fig. 5 are localized in the precuneus. In addition to the precuneus dominated ICs, other ICs with strong alpha peaks in their frequency spectra are recovered by each ICA method, including mu rhythm generators along the central sulcus and in posterior parietal areas. Fig. 6 displays other alpha-dominated ICA components. Figs. 6a and b show, respectively, the time course and MFT maps for the most representative alpha IC from each method, localized primarily on the left superior parietal area. It is clear that even when just one slice of data is displayed some components show focal activity in more than one place. There are many ways that ICA analysis in general, and LIANA in particular, can be used to study alpha oscillations as they are modulated by sensory or motor tasks, but further discussion is beyond the scope of this paper.

Construction of the simulated somatosensory data embedded in MEG data

The simulation data were constructed by adding "computer-generated somatosensory activation" (CGSA) data to the resting data. The background activity was provided by a 1 min (124,980 sample points) continuous recording of resting data already described in the last section. A CGSA was computed for a dipole source in SI using the expression for a current dipole in a spherical conductor (Sarvas, 1987). This CGSA was added to the background signal in the form of a spike-like (Gaussian) wavelet just after 30 ms of an artificially defined regular marker latency to mimic the standard SI response following median nerve stimulation. The exact location of CGSA can be seen in Fig. 7, the center of the SI ROI (blue circle). The maximum strength of the magnetic field generated by the CGSA dipole at its peak was half of one standard deviation above the mean of the resting data signal across all channels and latencies. The use of a CGSA ensures that a single focal generator is present in the data and the choices made ensured that the simulation data set resembled real data as much as possible. The strength of the CGSA was chosen so that it represented a relatively weak, regular, and independent known generator that produced an average CGSA MEG signal with peaks similar to those that would be produced by a standard median nerve stimulus.

Because the data length can affect the effectiveness of ICA analysis, we ran two simulations: The first with 199 CGSA trials (the whole data set), and the second with 50 CGSA trials (a quarter of the data; 15 s, 31250 sample points). The two simulations produce very similar results, so for simulated data we show only the results from the shorter simulation. Analysis of the longer data sets will be described later with the analysis of MEG data obtained during real somatosensory stimulation.

Analysis of simulated evoked responses embedded in the resting data

We note that activity is always present in left and right SI, including in the resting data measurement, as demonstrated by the analysis of the resting data in the last section. In agreement with this expectation, ICs with strong left and right SI activities were

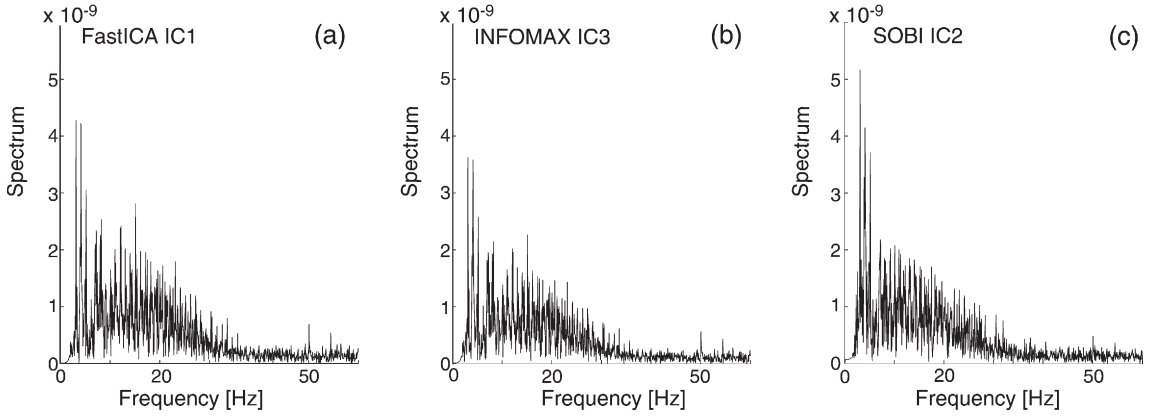

Fig. 4. FFT of the IC of each method that describes the cardiac signal.

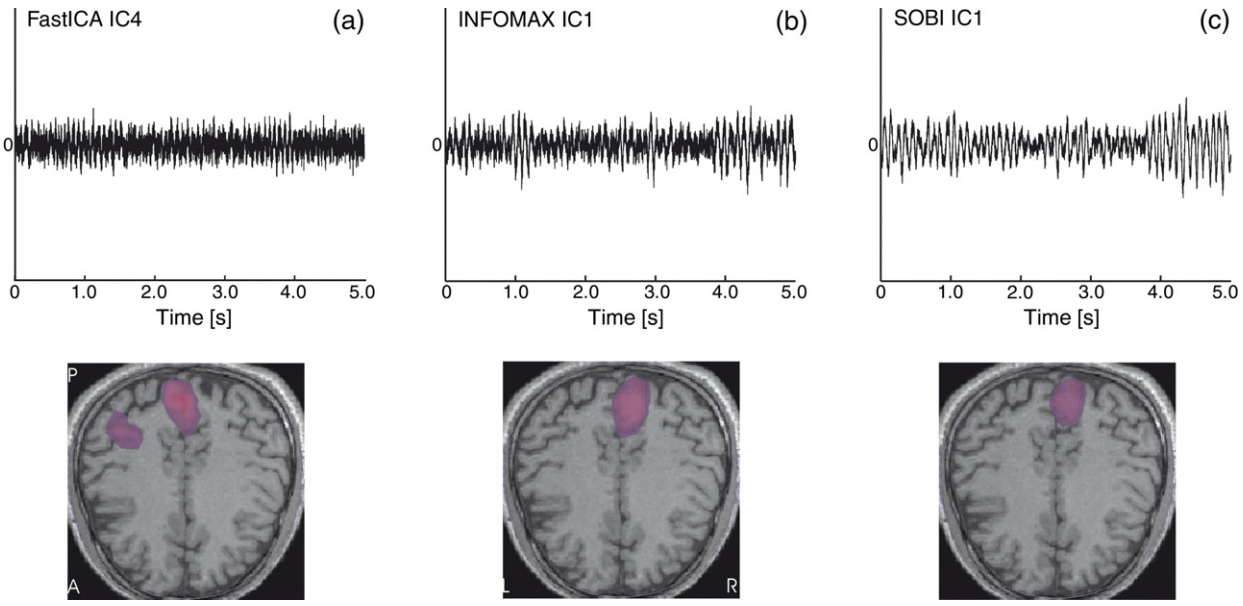

Fig. 5. The most representative alpha IC with strong precuneus focal activation. The time courses for FastICA (a), INFOMAX (b), and SOBI (c) are shown in the top rows and the activation maps in the bottom.

identified by ICA. We will compare the actual time course of the CGSA with the time courses of individual ICs, LIANA results, and full MFT solutions for activity in the left SI, where we have placed the CGSA. Fig. 7 shows the MFT reconstruction for each one of the selected ICs by LIANA. In each case, the map is superimposed onto an axial MRI slice of the subject's brain, cut through the maximum SI activity. Fig. 8 summarizes the results for the activation time course of the CGSA in the SI ROI. Since we combined CGSA and real resting data, the real amplitude changes in each location are not known. To show a better comparison of the activation curves, the computer-generated signal in Fig. 8 is normalized to the strongest peak in the recovered signal (SOBI LIANA) in Fig. 8b. Fig. 8a shows the results after averaging the 50 responses for the best IC (strongest value in ROI and/or lowest sort order and time-locked to stimulus onset) for each method. Other ICs were also identified with maxima outside SI but with relatively strong activity within the SI ROI (see Fig. 7). The final selection for LIANA was made from ICs that passed a given set threshold and showed time-locked behavior to the (artificial CGSA) sensory input. We have tested thresholds from 50% to 90%. The number of ICs differed considerably from method to method, especially at low thresholds. Fig. 8b shows the LIANA results with the

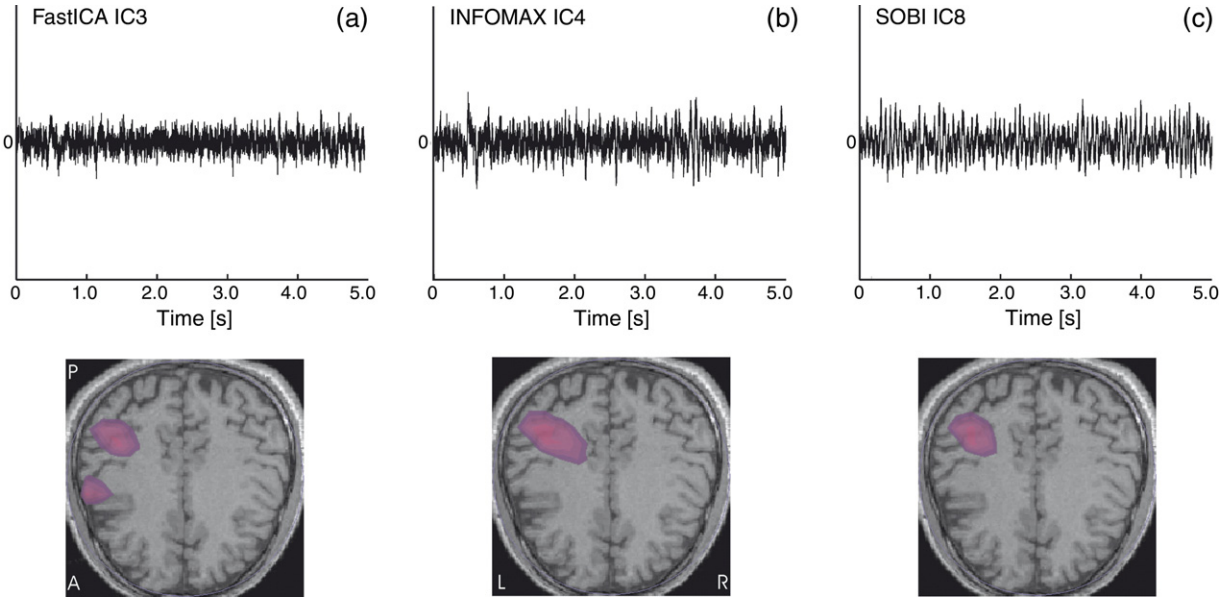

Fig. 6. The most representative alpha IC with strong left posterior parietal focal activation. The time courses for FastICA (a), INFOMAX (b), and SOBI (c) are shown in the top rows and the activation maps in the bottom.

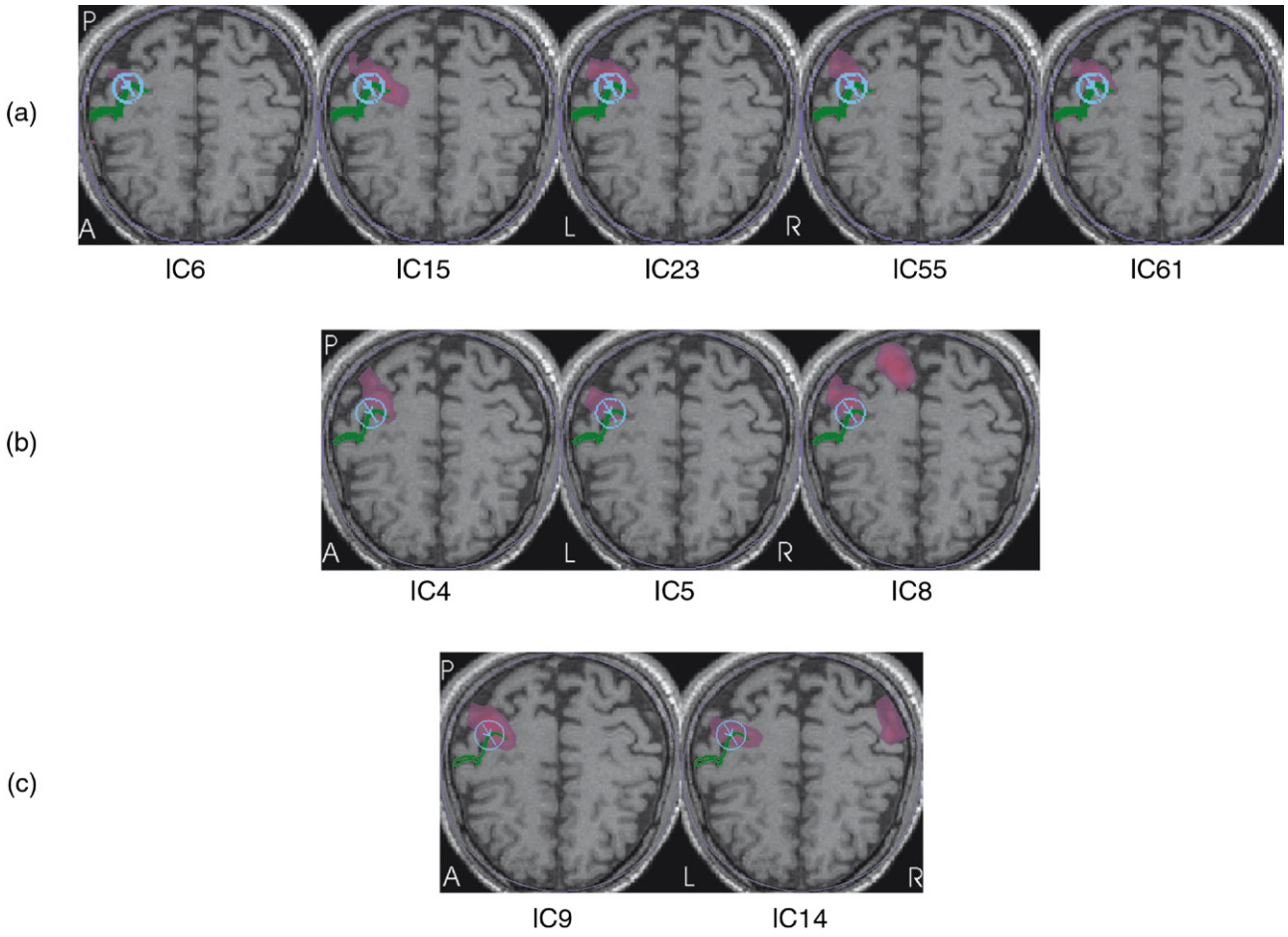

Fig. 7. The MFT solution for each IC selected by LIANA: (a) ICs selected for FastICA LIANA; (b) ICs selected for SOBI LIANA; and (c) ICs selected for INFOMAX LIANA. The green outline marks the central sulcus, and the circle, just behind the central sulcus, is the intersection of the spherical SI ROI with the displayed MRI slice. The peak current density of the MFT solution is at the center of the ROI and it is in the direction of the arrow (main direction). The CGSA was generated with a dipole placed at the center of the SI ROI and current dipole moment along the main direction.

following selections: two ICs (9, 14) for SOBI LIANA, three ICs (4, 5, 8) for INFOMAX LIANA, and 5 ICs (6, 15, 23, 55, 61) for FastICA LIANA, all using threshold 65%. Initially, two to ten ICs for these three methods passed the threshold but some were removed because they showed no time-locked response. For the best IC, IC9 for SOBI, IC5 for INFOMAX, and IC6 for FastICA were selected. The remaining panels of Fig. 6 show the results for two single trials at two different time scales. The plots in the middle row (Figs. 8c and d) are taken over the same time range as the average signal in the first row. In the bottom row, the same two trials (Figs. 8e and f) are displayed over a wider latency range. It appears that for the single trial shown in Figs. 8c and e, the CGSA is added onto a fairly quiet background, while for the single trial in Figs. 8d and f the CGSA is added onto a rather active background. Closer examination reveals that in both cases the evoked response is added to a slower activity, roughly in the alpha range. For the single trial in Figs. 8c and e, the CGSA is in phase, and for the single trial of Figs. 8d and f, in anti-phase with the alpha wave in the background. The results in Fig. 8 demonstrate that the single CGSA is separated into different ICs and since some of the ICs with fairly strong activity within SI have an overall maximum outside the SI ROI they must necessarily contain contributions from different brain generators. The separation of a single generator contribution into many ICs, as well as the mixing of contributions from different generators, varies from one ICA method to the next, as would be expected, given the different assumptions about the probability density function for the generators made by each method. Despite differences in the best IC, LIANA produces very similar activation within the targeted ROI for each ICA method by combining the contributions from all selected ICs.

Analysis of somatosensory evoked responses

LIANA and full MFT analysis were then applied to MEG data recorded during stimulation of the median nerve. The data were recorded at a high sampling rate (2 kHz) producing well over half a million (528041) sample points in one 4 min run. It is not possible to run ICA on the entire set at once. Therefore, we have split the 4 min run into three segments and performed a separate analysis on each. We have also down-sampled the data and run the analysis again as we will describe below. We will refer hereafter to one such segment as an epoch. We compared the LIANA results derived from the first epoch (first 1/3 of data, 174972 sample points), with the corresponding results obtained after low-pass filtering at 200 Hz and reduction of the data by 1/3

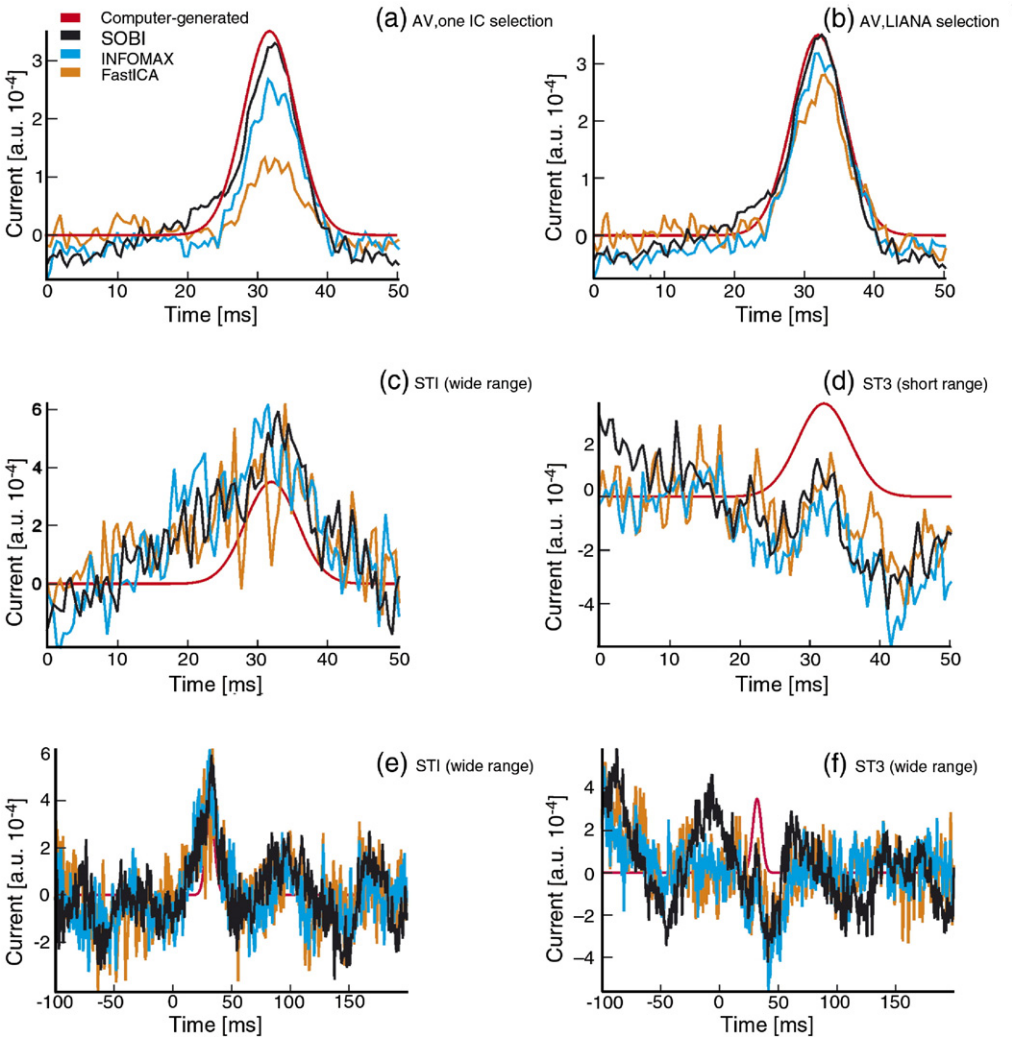

Fig. 8. Results of ICA and LIANA analysis of a mixture of background and computer-generated data. Panel a shows the average of 50 trials (one epoch of data) derived from the best IC for each ICA method. Panel b shows the same average (same 50 trials and ICA methods) for LIANA using all ICs that exceeded a threshold and showed a time-locked response. Panels c–f show LIANA estimates for two single trials. First, in panels c and d, the different LIANA estimates are shown in the latency range of the evoked response (0 ms–50 ms). In panels e and f, the responses are shown over the wider range (−100 ms–200 ms).

(to 176013 sample points). All ICA algorithms produced similar results for these two data sets, so we will only show results for the first data set containing 43 stimuli. We show results for two ROIs in the left hemisphere contralateral to the stimulated side. The first area is the SI and the second ROI is a little over 2 centimeters more posterior, situated in the SPL. Figs. 9 and 10 compare the frequency spectra and time courses of the activity in these two ROIs as computed with MFT (using all ICs not contaminated by artifacts) and with LIANA for each of the three ICA methods. The LIANA computations in the SI ROI used two ICs (10, 14) for SOBI, three ICs (7, 52, 65) for INFOMAX, and three ICs (7, 28, 44) for FastICA. In the SPL ROI, five ICs (2, 10, 14, 21, 32) were used for SOBI, six ICs (1, 7, 10, 11, 16, 40) for INFOMAX, and three ICs (11, 13, 23) for FastICA.

For the 43 trials with stimulation to the right wrist in one epoch, Fig. 9 shows the fast Fourier transform (FFT) of the concatenated, chopped data (from 100 ms before, to 200 ms after stimulus onset, i.e., 625 samples). The spectra were computed from the activation curves for the left SI (Fig. 9a) and left SPL (Fig. 9b) ROIs. Peaks at 3.33 Hz and 6.67 Hz, corresponding to the trial repetition rate, can be seen in the spectra of both MFT and the LIANA sets, and with each ICA method. As expected these peaks are more evident in the SI than SPL spectra. We therefore conclude that both MFT and LIANA, computed using ICs from any one of the three ICA methods, recover the dominant trial-by-trial response to each stimulus well.

Fig. 10 shows the time courses for the same two areas and types of computation. For the LIANA results, we compared the activation obtained by the combination of ICs satisfying the selection criteria with the activation for just the strongest one IC. For SI, the time course for this one IC component (IC10 for SOBI, IC7 for INFOMAX and FastICA) was almost indistinguishable from the LIANA solution for each one of the ICA algorithms. The SI activity was identified in two main waves, the first close to 20 ms and the second just after 50 ms. The two waves were evident in all methods. For the SPL area, the best-evoked response was with IC21 for SOBI, IC7 for INFOMAX and IC13 for FastICA. Although the activation time course for each one of these “best

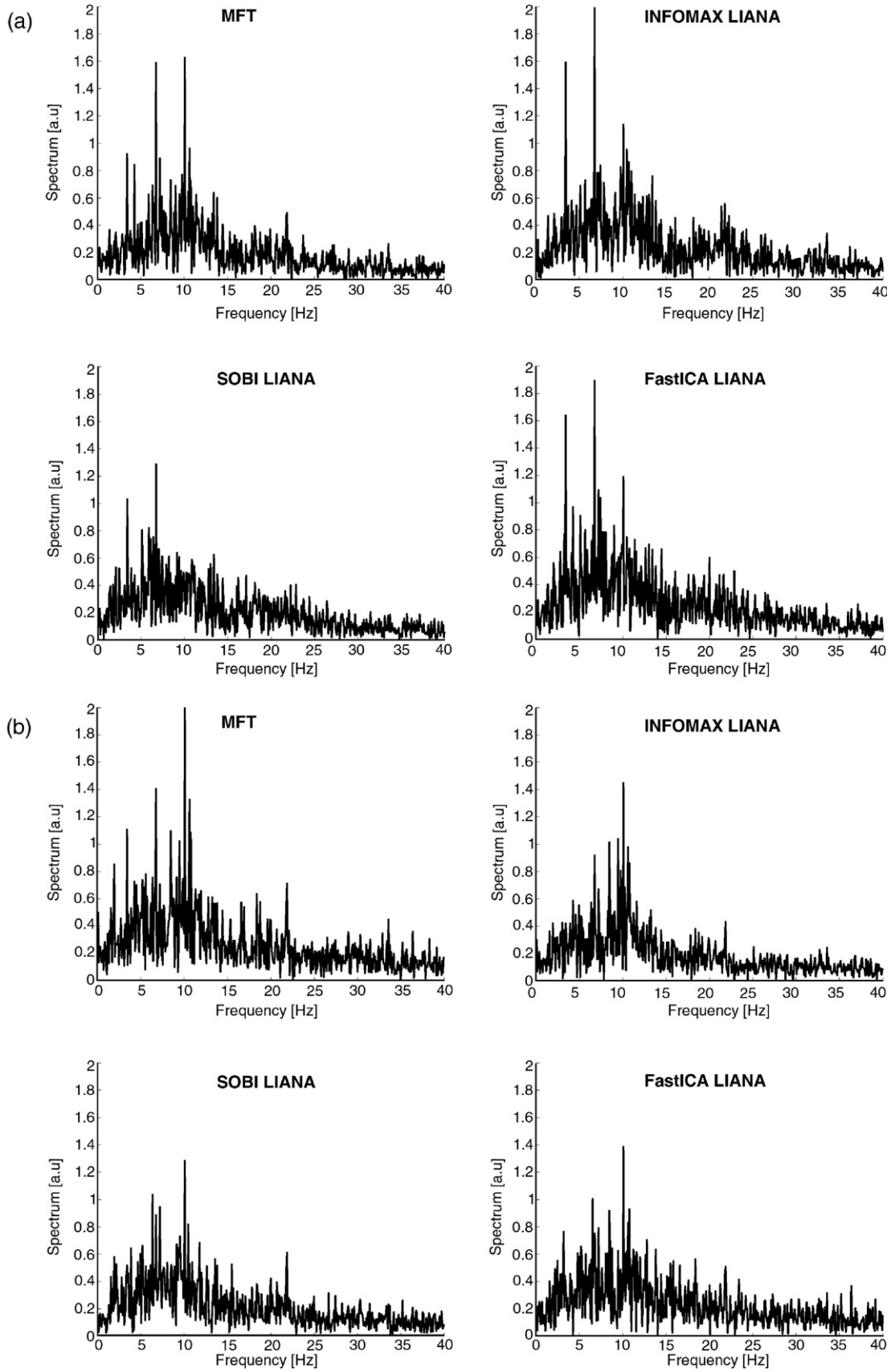

Fig. 9. Power spectrum of concatenated chopped activation curves for SI left (a) and SPL left (b) obtained from the same data set by applying the same ICA algorithms. All figures show strong 6.67 Hz peaks. The 3.33 Hz peaks are evident in most figures but are somewhat weaker for INFOMAX LIANA and SOBI LIANA in panel b.

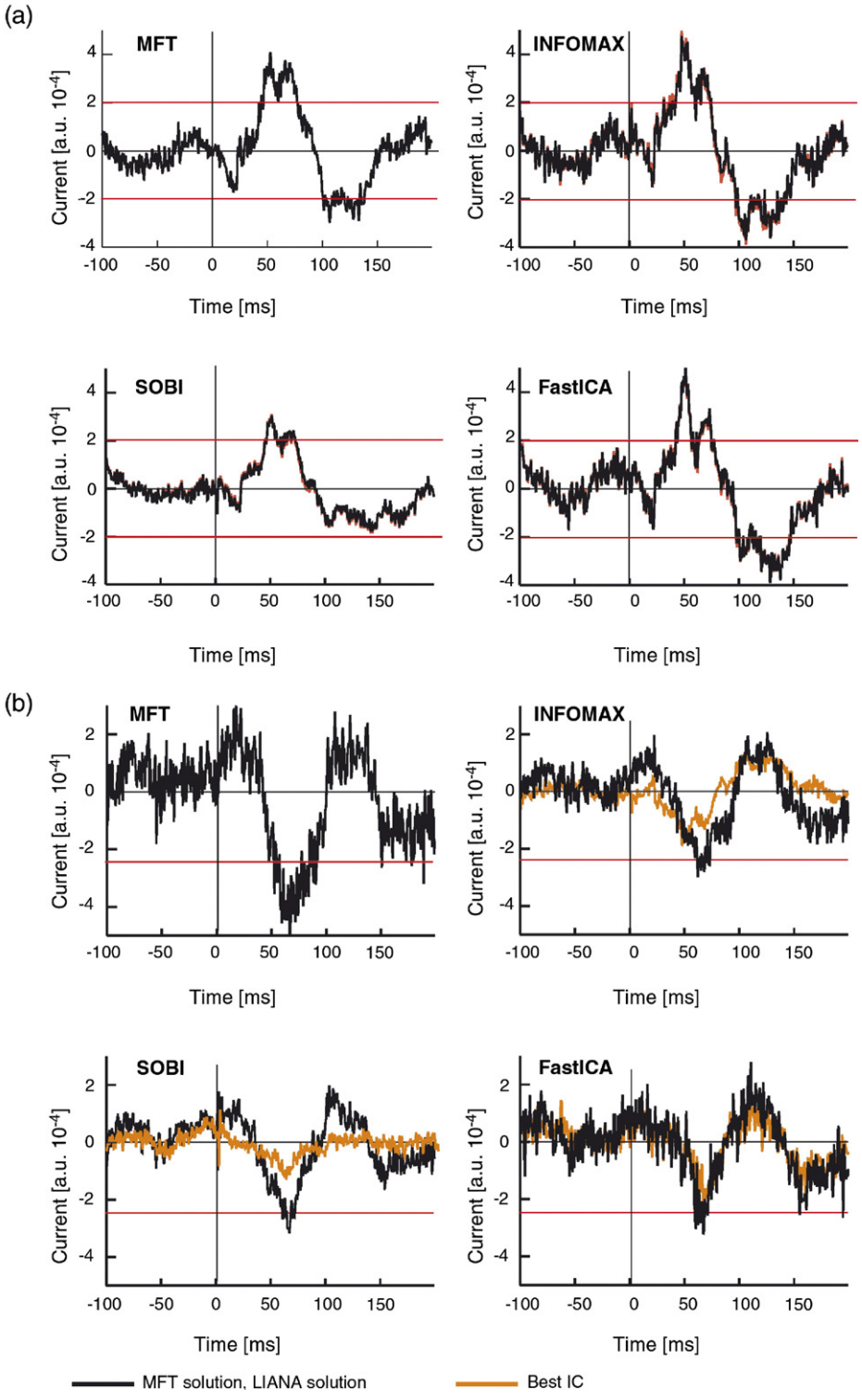

Fig. 10. Activation curves for regions of interest in the left SI (a) and left SPL (b). In each case, the activation curves are from the full MFT solution, using LIANA from the subspace defined by ICs (black curves), and the best IC (orange curve) for INFOMAX, SOBI, and FastICA. In the case of SI, Best IC adequately describes the response for all ICA methods tested.

ICs" was different the use of all ICs satisfying the LIANA selection criteria successfully recovered the activation, producing very similar results for the full MFT and each ICA method, showing a peak around 70 ms in each case.

Single-trial analysis

Table 1 shows the detection level for LIANA using the three ICA methods and full MFT single-trial analysis. The

Table 1
Detection level achieved by LIANA from a single ICA run containing in total 43 trials

	First SI peak		Second SI peak		SPL peak	
	Number of detected trials	Detection level %	Number of detected trials	Detection level (%)	Number of detected trials	Detection level (%)
SOBI LIANA	21	48.8	37	86.0	37	86.0
INFOMAX LIANA	25	58.1	38	88.4	38	88.4
FastICA LIANA	27	62.8	40	93.0	36	83.7
MFT	21	48.8	39	90.7	42	97.7

The current strength threshold was set to the level marked by the red line in Fig. 10 and the latency range was from 20.0 to 23.0 ms for the first and from 50.0 to 53.0 ms for the second SI peaks. For the SPL, the latency range was from 65.0 to 69.0 ms.

rather high threshold used for the detection is indicated by the red lines in Fig. 10. The definition of detection is defined as:

$$||f(a \le t \le b)|| \ge ||\text{threshold}||, \quad (13)$$

where f (=ACT_{ROI} (t) in Eq. (11)) is the time course activation function; t means time course; and [a, b] means target latency. For the first SI activity, threshold=$2.0*10^{-4}$, and [a, b]=[20.0, 23.0] ms (see Fig. 10a); for second SI activity, threshold=$-2.0*10^{-4}$, and [a, b]=[50.0, 53.0] ms. For the SPL, threshold=$-2.5*10^{-4}$, and [a, b]=[65.0, 69.0] ms. All four methods have similar detection ability especially for the second SI peak (see Fig. 10b). The higher detection level for INFOMAX and FastICA for the first SI peak is due to the mixing of background activity with the evoke responses. Table 2 lists which trials were missed at the second SI peak for each method. Most of the trials were missed by three or more methods and only rarely were isolated trials missed by just one method. Table 3 lists which trials were missed at the left SPL peak by each method. Two trials (22, 38) were missed by all LIANA methods.

Discussion

ICA has become a popular tool for MEG/EEG analysis, either as a tool for identifying and eliminating unwanted components, or for exploring further activations of interest. Most studies focus on identifying one IC. The implicit assumption is that ICA recovers the original signals with ICs that relate one-to-one to an underlying generator (see also Fig. 1). The model behind popular methods of MEG/EEG analysis, including the use of single ICs as descriptors of regional activations, is that it is possible to use some process "x" that can produce the "ideal output" column of Fig. 1. It is possible to extract individual ICs that show interesting behavior, or as in the case of SI in our example, a nearly complete activation. This is only achieved for activities that truly dominate the signal and/or have statistical properties emphasized by the particular ICA method used. It is not possible for all regional activations to dominate the signal and it is unlikely that the activity in different brain regions will share the same statistical properties. There is no guarantee, therefore, that any one, let alone all, of the identified ICs will be describing just one generator. This becomes immediately apparent if one uses the same ICA method on successive runs of the same data. The examples we have already described clearly demonstrate this mixing of components. The advantage of ICA is that it can perform remarkably well as a blind process. This is also its major disadvantage; the results of a blind method are not easy to generalize. We now propose a new method, LIANA, that follows from the realization that any one focal source may appear together with other sources in one, or more often, a few ICs, and that any one IC is usually a mixture of underlying generators. The basic steps for LIANA are summarized in Fig. 2. LIANA removes the blind aspect of the ICA process by focusing on just one or a few ROIs. The ROIs can be defined on anatomical grounds or from separate analysis (e.g., using conventional localization of the average signal). In summary, we use ICA to derive candidate components that are more likely to represent aspects of true generators than the raw MEG signals do. We can emphasize whatever measure is more appropriate for our problem by the choice of ICA algorithm, but this choice only sharpens the results by separating out the effect of interest in one or a few components. As summarized in Fig. 2, tomographic analysis of each IC allows direct identification of the subset of ICs that localize in one or more ROIs. The net activation curve for the set of these components is reconstructed using the tomographic description of each of the identified ICs and its ICA-derived time course.

The identification and removal of large artifacts is part of the ICA, and as Figs. 3 and 4 show, all ICA methods do well in the current investigation, at least for the cardiac response. The results in Figs. 3 and 4 suggest that FastICA and INFOMAX are best suited for spike signal detection, while SOBI produces the best result for sources with a well-defined, often oscillatory time course. However, one or another ICA algorithm is likely to best detect some of the many types of MEG/EEG artifacts (blinks, strong cardiac, and power line noise). Complete artifact elimination requires selection of one or more ICA algorithms, depending on the nature of the expected artifacts.

Table 2
The index number of single trials that failed to produce a second SI peak above the threshold (red line in Fig. 10a) in the latency range from 50.0 to 53.0 ms

SOBI	LIANA	1,4,16,18, 24, 38
INFOMAX	LIANA	4,16, 24, 37,38
FastICA	LIANA	4, 24, 36
MFT		4, 24, 36,37

Table 3
The index number of single trials that failed to produce an SPL peak above the threshold (red line in Fig. 10b) in the latency range from 65.0 to 69.0 ms

SOBI	LIANA	38
INFOMAX	LIANA	6, 12, 22, 27, 38
FastICA	LIANA	5, 7, 22,23,25, 38,43
MFT		13, 15,18,22, 29,38

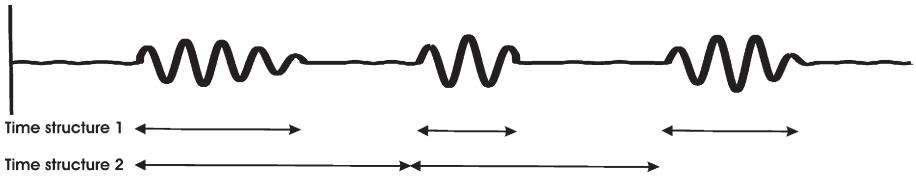

Fig. 11. Schematic representation of brain oscillatory activity. Time structure 1: Brain oscillatory activations are distinct by frequency range (alpha, beta, gamma, theta, etc.). Time structure 2: Periodic structure from one activation to the next.

Two conclusions can be drawn from the results displayed in Figs. 7 and 8. First, more than one components can contribute to the activation in a given ROI (Fig. 7). LIANA is able to recover the full ROI activation by combining all ICs that have strong activity in that ROI. Even for the specific case of the rather weak but consistent CGSA generator, practically any one ICA algorithm can identify the CGSA using LIANA. Second, simply using LIANA to sum the contributions from different ICs mixes together different types of responses. In the example we have provided, the alpha activity from the background test signal is mixed together with the "evoked" CGSA response, as is evident in Figs. 8e–f. For this case, the CGSA was truly independent since it was just added many times in exactly the same form to the background signal. For this artificial situation, the background oscillations were practically eliminated by averaging the 50 single trials aligned to the "onset" of stimulation. Our analysis with computer-generated data (Fig. 8) and real data (Figs. 9 and 10) showed that while different ICA methods produce different individual components, LIANA gives essentially the same results when it is applied to any one of them (particularly, Figs. 8a, b, and 10b). So, for the ROIs examined, LIANA provided the regional brain activation as effectively as a full single-trial MFT analysis (Figs. 9 and 10, Tables 1–3), but at a fraction of the computational cost. The computation is sufficiently fast to be competitive even for the case where just one IC adequately describes the activity, as was the case for the SI response in the LIANA result (Fig. 10a).

In conclusion, we have introduced a method, LIANA, for extracting regional activations from a subspace spanned by ICs, namely, those ICs that localize within the ROI. Other criteria can also be easily incorporated, e.g., time locking to channels monitoring internal events or external stimuli. We have demonstrated that LIANA can successfully reconstruct the activations equally well when a single IC dominates and when the signal generated by the targeted, underlying generator is spread over multiple ICs.

Acknowledgments

Special thanks to Allan K. Barros for theoretical discussions, Daniel Palomo for general comments and redaction, Max Burbank, Kenji Haruhana, Masahiro Machizawa for MEG data acquisition, Vahe Poghosyan for discussions and help with data analysis and computing, and Peter Fenwick and Afiza Abu-Bakar for general comments on an earlier version of the manuscript.

Appendix A

We introduce in this appendix the ICA algorithms employed in this paper and whenever necessary list the implementation parameters we have used.

A.1. ICA algorithms

A.1.1. INFOMAX

The algorithm relies on the factorization of marginal distributions,

$$p(y_1,\cdots,y_n) = p(y_1)\cdots p(y_n), \tag{A.1}$$

where p means probability density function. By minimizing the Kulback–Leibler information it follows,

$$\Delta W \propto -\eta(I-\varphi(y)y^T)W, \tag{A.2}$$

where η is a learning rate parameter, T denotes transpose operation. This algorithm works well for super-Gaussian type signals; however, it is not suitable for sub-Gaussian distributions. Therefore, the following algorithm was proposed.

A.1.2. INFOMAX(-EXT)

The algorithm was generalized to,

$$\Delta W \propto \eta(I-K\tanh(y)y^T - yy^T)W, \tag{A.3}$$

where K is assigned according to an estimate of the kurtosis of the data. For super-Gaussian distributions K is set to 1, while for sub-Gaussian distributions K is set to -1.

A.1.3. FastICA (using kurtosis)

The basic concept of this algorithm is that non-Gaussian components are independent. Since kurtosis is zero for a Gaussian, maximizing or minimizing the kurtosis under the constraint, the local extrema is equivalent to estimating a non-Gaussian independent component. In FastICA algorithm, we assume input signal is sphered. Considering a linear combination of x using unit vector w, to maximize or minimize $kurt(w^Tx)$, the following iteration can be used,

$$\Delta w \propto E\{x(w^Tx)^3\} - 3w, \tag{A.4}$$

where E denotes expectation. However, this algorithm can be sensitive to outliers, hence the following algorithm was proposed.

A.1.4. FastICA (using negentropy)

Negentropy is a normalized version of differential entropy to obtain a measure of non-Gaussianity, it is also zero for Gaussian distributions.

$$\Delta w \propto E\{xg(w^Tx)\} - E\{g'(w^Tx)\}w \tag{A.5}$$

Eq. (A.4) is obtained by solving a Lagrange multiplier using Newton's method. For the transformation function, $g()$, a cubic

function is used for super-Gaussian and hyperbolic tangent function for sub-Gaussian as defaults.

A.1.5. SOBI

This algorithm takes into an account assumed time structure of the input signal by the joint diagonalization of correlation matrixes,

$$M = E\{x(t+\alpha)x(t)^T\}, \quad (A.6)$$

$$D = JMJ^T, \quad (A.7)$$

where x is (whitened) input signals, α is time-delayed parameter, and J is the Jacob angle matrix.

A.2. Implementation parameters

All ICA algorithms employed in this paper have some free parameters that must be adjusted to fit the requirements of the problem. The performance of the various algorithms depends on the level of noise that may also affect the optimal choice for the free parameter values. Clearly an exhaustive study of all algorithms for different types of signals and under different noise levels is not feasible. In order to maintain consistency with earlier studies, we used source codes from Web sites (FastICA [1*], INFOMAX (EEGLAB[2*]), JADE [3*]). The complete code for SOBI was not available; however, a correlation matrix using time-delayed parameters is easy to code and the joint diagonalization parts followed the approach used in JADE. All the code was originally available in MATLAB and was translated to IDL code to allow direct integration with our in-house MEG analysis and localization software. Accuracy tests were performed to pre-defined accuracy using double precision floating point. Whenever possible, we selected default parameters. We used as defaults INFOMAX-EXT (Eq. (A.3)) and FastICA using negentropy (Eq. (A.5)). The final choices required tuning parameters for INFOMAX and the time-delayed parameters for SOBI. Some tuning parameters for INFOMAX, i.e., learning rate, block size, kurt-size, are required for optimization. For INFOMAX-EXT for example, kurtosis was estimated using all points at first, but whenever this was not practical it was reduced to less than 10000 points. If INFOMAX-EXT did not convergence after long computation, i.e., 1 day, after some parameter adjustment, we used instead INFOMAX (Eq. (A.2)).

We now discuss SOBI in some detail because in agreement with other recent studies, we found that it effectively exploits regularities in the signal time structure. The commercial power line (50 Hz in our case) and its harmonics are typical noise with time structure and are present during an entire recording period. Cardiac function is also a constant noise for MEG measurement with default time structure which changes slightly over times. The main components with time structure from brain activities are oscillatory rhythms. Previous reports indicate that brain rhythmic signals generated from different sources usually have their own oscillatory frequencies with distinct phases and are located in specific brain regions with patterns that are distinct from artifacts and noise. Fig. 11 shows a prototyped structure of oscillatory activity and the relevant time scales. These time scales are defined by the intrinsic frequency of the oscillations and by the repetition rate of oscillatory bursts. The objective of SOBI algorithm is to maximize the auto-correlation with the time-delayed parameter and minimize the cross-correlation. Both time structures 1 and 2 in Fig. 11 are important, and the second one becomes increasingly important for long recording. Unfortunately time structure 2 is not usually found and it often changes over time. Small time delays are useful for detecting sudden events like outliers. In summary for the sampling rate of 2083 Hz, which is higher than usual for MEG/EEG recoding, the time-delayed parameters with cardiac (2083 (1 Hz)), power line (40 (50 Hz) and its multiples), oscillatory brain activations (416 (5 Hz), 208 (10 Hz), … 30 (70 Hz) and these multiples), and small time delay (1, 2, 3,…, 25) were selected.

time lag=[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 104, 208, 312, 416, 520, 728, 832, 936, 1040, 1560, 2083]

[1*], http://www.cis.hut.fi/projrcts/ica/fastica, Version 2.1, January 15 2001

[2*], http://www.sccn.ucsd.edu/~scott/ica.html, eeglab4.512

[3*], http://www.tsi.enst.fr/~cardoso/giodesepsou.html, Version 1.5 (November 2, 1997)

References

Amari, S., 1998. Natural gradient works efficiently in learning. Neural Comput. 10, 251–276.

Barbati, G., Porcaro, C., Zappasodi, F., Rossinia, P.M., Tecchio, F., 2004. Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals. Clin. Neurophysiol. 115, 1220–1232.

Barros, A.K., Mansour, A., Ohnishi, N., 1998. Removing artifacts from electrocardiographic signals using independent components analysis. Neurocomputing 22, 173–186.

Bell, A.J., Sejnowski, T.J., 1995. An information maximization approach to blind separation and blind deconvolution. Neural Comput. 7, 1129–1159.

Belouchrani, A., Abed-Meraim, K., Cardoso, J.F., Moulines, E., 1997. A blind source separation technique using second-order statistics. IEEE Trans. Signal Process. 45, 434–444.

Cardoso, J.F., Souloumiac, A., 1996. Jacobi angles for simultaneous diagonalization. SIAM J. Matrix Anal. Appl. 17, 161–164.

Comon, P., 1994. Independent component analysis, a new concept? Signal Process. 36, 287–314.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D.Y., Engel, A.K., 2005. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. J. Neurosci. 25, 11730–11737.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.

Delorme, A., Makeig, S., Fabre-Thorpe, M., Sejnowski, T.J., 2002. From single-trial EEG to brain area dynamics. Neurocomputing 44–46, 1057–1064.

Hamalainen, M., Hari, R., Ilmoniemi, R.J., Knuutila, J., Lounasmaa, O.V., 1993. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev. Modern Phys. 65, 413–497.

Hironaga, N., Schellens, M., Ioannides, A.A., 2002. Accurate co-registration for MEG reconstructions, Hannes Nowak, Jens Haueisen, Frank GieBler, Ralph Huonker, VDE Verlag, Berlin. Proceedings of the 13th International Conference on Biomagnetism, pp. 931–933.

Hyvarinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans. Neural Netw. 10, 626–634.

Hyvarinen, A., Oja, E., 1997. A fast fixed-point algorithm for independent component analysis. Neural Comput. 9, 1483–1492.

Ioannides, A.A., 2001. Real time human brain function: observations and inferences from single trial analysis of magnetoencephalographic signals. Clin. Electroencephalogr. 32, 98–111.

Ioannides, A.A., Fenwick, P.B.C., 2005. Imaging cerebellum activity in real time with magnetoencephalography. Prog. Brain Res. 148, 139–150.

Ioannides, A.A., Taylor, J.G., 1999. Minimum norm, magnetic field tomography and FOCUSS. In: Yoshimoto, T., Kotani, M., Kuriki, S., Karibe, H., Nakasato, N. (Eds.), Recent Advances in Biomagnetism. Tohoku University Press, pp. 228–231.

Ioannides, A.A., Bolton, J.P.R., Clarke, C.J.S., 1990. Continuous probabilistic solutions to the biomagnetic inverse problem. Inverse Probl. 6, 523–542.

Ioannides, A.A., Corsi-cabrera, M., Fenwick, P.B.C., Portilla, Y.D., Laskaris, N.A., Khurshudyan, A., Theoflou, D., Shibata, T., Uchida, S., Nakabayashi, T., Kostopoulos, G.K., 2004. MEG tomography of human cortex and brainstem activity in waking and REM sleep saccades. Cereb. Cortex 14, 56–72.

Ioannides, A.A., Fenwick, P.B.C., Liu, L.C., 2005. Widely distributed magneto-encephalography spikes related to the planning and execution of human saccades. J. Neurosci. 25, 7950–7967.

Joyce, C.A., Gorodnitsky, I.F., Kutas, M., 2004. Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. Psychophysiology 41, 313–325.

Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., Mckeown, M.J., Iragui, V., Sejnowski, T.J., 2000a. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37, 163–178.

Jung, T.P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T.J., 2000b. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin. Neurophysiol. 111, 1745–1758.

Jung, T.P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T.J., 2001. Analysis and visualization of single trial event-related potentials. Hum. Brain Mapp. 14, 166–185.

Kishida, K., Fukai, H., Hara, T., Shinosaki, K., 2003. A new approach to blind system identification in MEG data. IEICE Trans. Fund. E86-A, 611–619.

Kobayashi, K., James, C.T., Nakahori, T., Akiyama, T., Gotman, J., 1999. Isolation of epileptiform discharges from unaveraged EEG by independent component analysis. Clin. Neurophysiol. 110, 1755–1763.

Kobayashi, K., Akiyama, T., Nakahori, T., Yoshinaga, H., Gotman, J., 2002a. Systematic source estimation of spikes by a combination of independent component analysis and RAP-MUSIC I: principles and simulation study. Clin. Neurophysiol. 113, 713–724.

Kobayashi, K., Akiyama, T., Nakahori, T., Yoshinaga, H., Gotman, J., 2002b. Systematic source estimation of spikes by a combination of independent component analysis and RAP-MUSIC II: preliminary clinical application. Clin. Neurophysiol. 113, 725–734.

Laskaris, N.A., Liu, L.C., Ioannides, A.A., 2003. Single-trial variability in early visual neuromagnetic responses: an explorative study based on the regional activation contributing to the N70m peak. NeuroImage 20, 765–783.

Lee, T.W., Girolami, M., Sejnowski, T.J., 1999. Independent component analysis using an extended INFOMAX algorithm for mixed sub-Gaussian and super-Gaussian sources. Neural Comput. 11, 606–633.

Lee, P.L., Wu, Y.T., Chen, L.F., Chen, Y.S., Cheng, C.M., Yeh, T.C., Ho, L. T., Chang, M.S., Hsieh, J.C., 2003. ICA-based spatiotemporal approach for single-trial analysis of postmovement MFG beta synchronization. NeuroImage 20, 2010–2030.

Liu, L.C., Ioannides, A.A., Muller-Gartner, H.W., 1998. Bi-hemispheric study of single trial MEG signals of the human auditory cortex. Electroencephalogr. Clin. Neurophysiol. 106, 64–78.

Makeig, S., Jung, T.P., Bell, A.J., Ghahremani, D., Sejnowski, T.J., 1997. Blind separation of auditory event-related brain responses into independent components. Proc. Natl. Acad. Sci. U. S. A. 94, 10979–10984.

Makeig, S., Westerfield, M., Jung, T.P., Covington, J., Townsend, J., Sejnowski, T.J., Courchesne, E., 1999. Functionally independent components of the late positive event-related potential during visual spatial attention. J. Neurosci. 19, 2665–2680.

Makeig, S., Westerfield, M., Jung, T.P., Enghoff, S., Townsend, J., Courchesne, E., Sejnowski, T.J., 2002. Dynamic brain sources of visual evoked responses. Science 295, 690–694.

Molgedey, L., Schuster, H.G., 1994. Separation of a mixture of independent signals using time delayed correlations. Phys. Rev. Lett. 72, 3634–3637.

Moradi, F., Liu, L.C., Cheng, K., Waggoner, R.A., Tanaka, K., Ioannides, A. A., 2003. Consistent and precise localization of brain activity in human primary visual cortex by MEG and fMRI. NeuroImage 18, 595–609.

Ossadtchi, A., Balliet, S., Mosher, J.C., Thyerlei, D., Sutherling, W., Leahy, R.M., 2004. Automated interictal spike detection and source localization in magnetoencephalography using independent components analysis and spatio-temporal clustering. Clin. Neurophysiol. 115, 508–522.

Sander, T.H., Burghoff, M., Curio, G., Trahms, L., 2005. Single evoked somatosensory MEG responses extracted by time delayed decorrelation. IEEE Trans. Signal Process. 53, 3384–3392.

Sarvas, J., 1987. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Phys. Med. Biol. 32, 11–22.

Scherg, M., 1990. fundamentals of dipole source potential analysis, in auditory evoked magnetic fields and electric potentials. Adv. Audiol. 6, 40–69.

Tang, A.C., Pearlmutter, B.A., Malaszenko, N.A., Phung, D.B., 2002a. Independent components of magnetoencephalography: single-trial response onset time. NeuroImage 17, 1773–1789.

Tang, A.C., Pearlmutter, B.A., Malaszenko, N.A., Phung, D.B., Reeb, B.C., 2002b. Independent components of magnetoencephalography: localization. Neural Comput. 14, 1827–1858.

Tang, A.C., Sutherland, M.T., McKinney, C.J., 2005. Validation of SOBI components from high-density EEG. NeuroImage 25, 539–553.

Taylor, J.G., Ioannides, A.A., Mueller-Gaertner, H.W., 1999. Mathematical analysis of lead field expansions. IEEE Trans. Med. Imag. 18, 151–163.

Vigario, R.N., 1997. Extraction of ocular artefacts from EEG using independent component analysis. Electroencephalogr. Clin. Neurophysiol. 103, 395–404.

Vigario, R., Sarela, J., Oja, E., 1998. Independent component analysis in wave decomposition of auditory evoked fields. Proceedings of the International Conference on Artificial Neural Netw. (ICANN'98).

Vigario, R., Sarela, J., Jousmaki, V., Oja, E., 1999. Independent component analysis in decomposition of auditory and somatosensory evoked fields. Proceedings of the International Workshop on Independent Component Analysis and Blind Separation of Signals (ICA'99).

Wallstrom, G.L., Kass, R.E., Miller, A., Cohn, J.F., Fox, N.A., 2004. Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods. Int. J. Psychophysiol. 53, 105–119.

Yeung, N., Bogacz, R., Holroyd, C.B., Cohen, J.D., 2004. Detection of synchronized oscillations in the electroencephalogram: an evaluation of methods. Psychophysiology 41, 822–832.