Imaging Brain Dynamics Using Independent Component Analysis

TZYY-PING JUNG, MEMBER, IEEE, SCOTT MAKEIG, MARTIN J. MCKEOWN, ANTHONY J. BELL, TE-WON LEE, MEMBER, IEEE, AND TERRENCE J. SEJNOWSKI, FELLOW, IEEE

Invited Paper

The analysis of electroencephalographic (EEG) and magnetoencephalographic (MEG) recordings is important both for basic brain research and for medical diagnosis and treatment. Independent component analysis (ICA) is an effective method for removing artifacts and separating sources of the brain signals from these recordings. A similar approach is proving useful for analyzing functional magnetic resonance brain imaging (fMRI) data. In this paper, we outline the assumptions underlying ICA and demonstrate its application to a variety of electrical and hemodynamic recordings from the human brain.

Keywords—*Blind source separation, EEG, fMRI, independent component analysis.*

I. INTRODUCTION

Independent component analysis (ICA) refers to a family of related algorithms [1]–[10] that exploit independence to perform blind source separation. In Section II, an ICA algorithm based on the Infomax principle [6] is briefly introduced. In Section III, ICA is applied to electroencephalographic (EEG) recordings. Although these weak signals recorded from the surface of the scalp have been studied for nearly 100 years, their origins, exact dynamics, and relationship to brain function has been difficult to assess

T.-P. Jung, S. Makeig, A. J. Bell, T.-W. Lee, and T. J. Sejnowski are with the University of California at San Diego, La Jolla, CA 92093-0523 USA and also with The Salk Institute for Biological Studies, La Jolla, CA 92037 USA (e-mail: jung@salk.edu; scott@salk.edu; tony@salk.edu; tewon@salk.edu; terry@salk.edu).

M. J. McKeown is with the Department of Medicine (Neurology), the Brain Imaging and Analysis Center (BIAC), and the Department of Biomedical Engineering, Duke University, Durham, NC 27708 USA (e-mail: martin.mckeown@duke.edu).

because signals recorded at the scalp are mixtures of signals from multiple brain generators. ICA may be helpful in identifying different types of generators of the EEG as well as its magnetic counterpart, the magnetoencephalogram (MEG). This application also illustrates questions concerning the assumptions required to apply ICA to biological time series. In Section IV, we show that ICA can also be used to analyze hemodynamic signals from the brain recorded using functional magnetic resonance imaging (fMRI). This exciting new area of research allows neuroscientists to noninvasively measure brain activity in humans indirectly through slower changes in brain blood flow. In all of these examples, great care must be taken to examine the validity of the assumptions that are used by ICA to derive a decomposition of the observed signals and/or to evaluate the reliability and functional significance of the resulting components.

II. INDEPENDENT COMPONENT ANALYSIS

ICA [4] was originally proposed to solve the blind source separation problem, to recover N source signals, $\mathbf{s} = \{s_1(t), \ldots, x_N(t)\}$, (e.g., different voice, music, or noise sources) after they are linearly mixed by multiplying by $\mathbf{A}$, an unknown matrix, $\mathbf{x} = \{x_{1(t)}, \ldots, x_N(t)\} = \mathbf{As}$, while assuming as little as possible about the natures of $\mathbf{A}$ or the component signals. Specifically, one tries to recover a version, $\mathbf{u} = \mathbf{Wx}$, of the original sources, $\mathbf{s}$, identical save for scaling and permutation, by finding a square matrix, $\mathbf{W}$, specifying spatial filters that linearly invert the mixing process. The key assumption used in ICA to solve this problem is that the time courses of activation of the sources (or in other cases the spatial weights) are as statistically independent as possible. Most ICA is performed using information-theoretic unsupervised learning algorithms. Despite its relatively short history, ICA is rapidly becoming a standard technique in multivariate analysis.

Mathematically, the ICA problem is as follows: We are given a collection of N-dimensional random vectors, $\mathbf{x}$

(sound pressure levels at N microphones, N-pixel patches of a larger image, outputs of N scalp electrodes recording brain potentials, or nearly any other kind of multidimensional signal). Typically there are diffuse and complex patterns of correlation between the elements of the vectors. ICA, like principal component analysis (PCA), is a method to remove those correlations by multiplying the data by a matrix as follows:

$$\mathbf{u} = \mathbf{Wx}. \tag{1}$$

(Here, we imagine the data is zero-mean; see below for pre-processing details.) But while PCA only uses second-order statistics (the data covariance matrix), ICA uses statistics of all orders and pursues a more ambitious objective. While PCA simply decorrelates the outputs (using an orthogonal matrix $\mathbf{W}$), ICA attempts to make the outputs statistically independent, while placing no constraints on the matrix $\mathbf{W}$. Statistical independence means the joint probability density function (pdf) of the output *factorizes*

$$p(\mathbf{u}) = \prod_{i=1}^{N} p_i(u_i) \tag{2}$$

while decorrelation means only that $\langle \mathbf{uu}^T \rangle$, the covariance matrix of $\mathbf{u}$, is diagonal (here, $\langle \rangle$ means average).

Another way to think of the transform in (1) is as

$$\mathbf{x} = \mathbf{W}^{-1}\mathbf{u}. \tag{3}$$

Here, $\mathbf{x}$ is considered the linear superposition of *basis functions* (columns of $\mathbf{W^{-1}}$), each of which is activated by an independent component, $\mathbf{u_i}$. We call the rows of $\mathbf{W}$ *filters* because they extract the independent components. In orthogonal transforms such as PCA, the Fourier transform and many wavelet transforms, the basis functions and filters are the same (because $\mathbf{W^T} = \mathbf{W^{-1}}$), but in ICA they are different.

The usefulness of a nonorthogonal transform sensitive to higher order statistics can be seen in Fig. 1, which shows the PCA and ICA basis functions for a simulated two-dimensional (2-D) non-Gaussian data distribution. Clearly the ICA axes capture much more about the structure of these data than the PCA. Similar data distributions are actually more common in natural data than those who model data by "mixtures of Gaussians" might suppose. This fact arises from the common nonorthogonal "mixing together" of highly sparse independent components. By sparse, we typically mean a distribution that is much "peakier" (e.g., near zero) than a Gaussian distribution, and with longer tails. A more technical term for sparse is super-Gaussian, usually identified with positive kurtosis.

The ICA problem was introduced by Herault and Jutten [1]. The results of their algorithm were poorly understood and led to Comon's 1994 paper defining the problem, and to his solution using fourth-order statistics. Much work took place in this period in the French signal processing community, including Pham *et al.*'s [3] Maximum Likelihood approach that subsequently formed the basis of Cardoso and
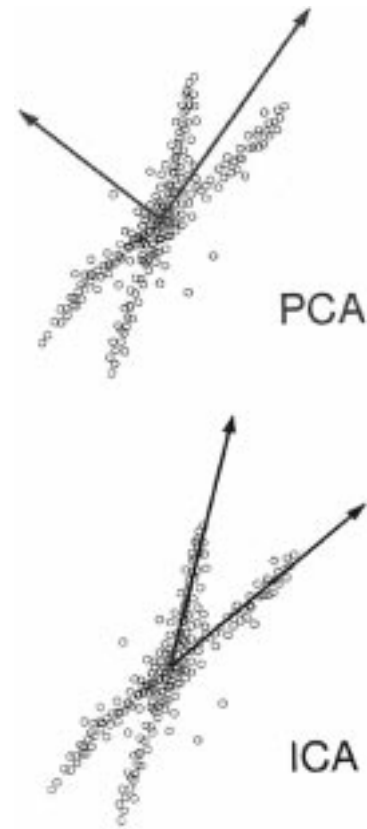

Fig. 1. The difference between PCA and ICA on a nonorthogonal mixture of two distributions that are independent and highly sparse (peaked with long tails). An example of a sparse distribution is the Laplacian: $\mathbf{p}(\mathbf{x}) = \mathbf{k}e^{|-\mathbf{x}|}$. PCA, looking for orthogonal axes ranked in terms of maximum variance, completely misses the structure of the data. Although these distributions may look strange, they are quite common in natural data.

Laheld's [7] EASI method. These methods are very close to the "Infomax" approach [6], so this algorithm may be called Infomax/ML ICA. Earlier, Cichocki *et al.* [5] had proposed an algorithm which motivated Amari [8] and colleagues to show that its success was due to its relation to a "natural gradient" modification of the Infomax/ML ICA gradient. This modification greatly simplified the algorithm, and made convergence faster and more stable.

The resulting gradient-descent algorithm (implemented for routine use by Makeig *et al.* (http://www.cnl.salk.edu/~scott/ica.html [11]) has proved useful in a wide range of biomedical applications. Batch algorithms for ICA also exist, such as Hyvärinen's *FastICA* and several cumulant-based techniques, including Cardoso's widely used fourth-order algorithm *JADE*. When these well-known algorithms are compared, they generally perform near equally well. However, applied to actual data sets for which no ground truth solutions exist, and for which the exactness of the ICA assumptions cannot be tested, they may produce differences whose relative value and significance are difficult to evaluate. Review papers comparing different ICA algorithms and their interrelationships are available [12], [13], as are two edited collections of papers [14], [15] and proceedings from two international workshops (ICA99, ICA2000). A third workshop in this series is planned (see
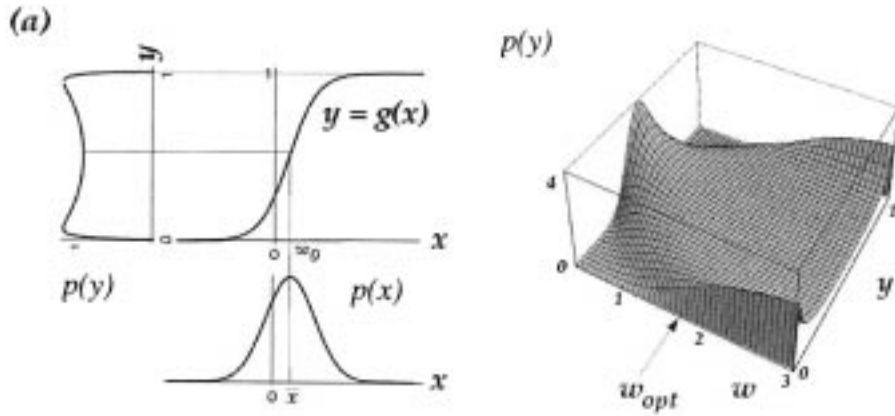
Fig. 2. Optimal information flow in sigmoidal neurons. (*left*) Input **x** raving probability density function $p(x)$, n this case a Gaussian, is passed through a nonlinear function $g(x)$. The information in the resulting density, $p(x)$ depends on matching the mean and variance of **x** to the threshold, $\mathbf{w_0}$, and slope, **w**, of $g(x)$ (Nicol Schraudolph, personal communication). (*right*) $p(y)$ is plotted for different values of the weight **w**. The optimal weight, $\mathbf{w_{opt}}$ transmits most information (from [2] by permission.)

http://ica2001.org). Matlab code for several algorithms, including those mentioned above, is also available through the World Wide Web. Below, we sketch the derivation and development of Infomax ICA.

A. The Infomax ICA Algorithm

A more general linear transform of **u** is the affine transform: $\mathbf{u} = \mathbf{Wx} + \mathbf{w}$ where **w** is an N-by-1 "bias" vector that centers the data on the origin. If we assume the independent component pdfs, $p_i(u_i)$ are roughly symmetrical, then it is simpler to subtract the mean, $\langle \mathbf{x} \rangle$, from the data beforehand. A second preprocessing step that speeds convergence is to first "sphere" the data by diagonalizing its covariance matrix

$$\mathbf{x} \leftarrow 2\langle \mathbf{xx}^T \rangle^{-1/2}(\mathbf{x} - \langle \mathbf{x} \rangle). \tag{4}$$

This yields a decorrelated data ensemble whose covariance matrix satisfies $\langle \mathbf{xx^T} \rangle = 4\mathbf{I}$, where **I** is the identity matrix. This is a useful starting point for ICA decomposition. This sphering method is not PCA, but rather zero-phase whitening which constrains the matrix **W** to be symmetric. By contrast, PCA constrains it to be orthogonal, and ICA, also a decorrelation technique but without constraints on **W**, finds its constraints in the higher order statistics of the data.

The objective of the Infomax ICA algorithm is to minimize *redundancy* between the outputs. This is a generalization of the mutual information

$$I(\mathbf{u}) = \int p(\mathbf{u}) \log \frac{p(\mathbf{u})}{\prod_{i=1}^{N} p_i(u_i)} d\mathbf{u}. \tag{5}$$

This redundancy measure has value 0 when the pdf **p(u)** factorizes, as in (2), and is a difficult function to minimize directly. The insight that led to the Infomax ICA algorithm was

that $\mathbf{I}(\mathbf{u})$ is related to the joint entropy, $\mathbf{H}(\mathbf{g}(\mathbf{u}))$, of the outputs passed through a set of sigmoidal nonlinear functions, g_i

$$I(\mathbf{u}) = -H(\mathbf{g}(\mathbf{u})) + E\left[\sum_i \log \frac{|g_i'(u_i)|}{p_i(u_i)}\right]. \tag{6}$$

Thus, if the absolute values of the slopes of the sigmoid functions, $|g_i'(u_i)|$ are the same as the independent component pdf's, $p_i(u_i)$ then Infomax [maximizing the joint entropy of the $\mathbf{g}(\mathbf{u})$ vector], will be the same as ICA (minimizing the redundancy in the **u** vector).

The principle of "matching" the $g_i's$ to the $p_i's$ is illustrated in Fig. 2, where a single Infomax unit attempts to match an input Gaussian distribution to a logistic sigmoid unit, for which

$$g(u) = \frac{1}{1 + e^{-u}}. \tag{7}$$

The match cannot be perfect, but it does approach the maximum entropy pdf for the unit distribution by maximizing the expected log slope, $E[\log|g'(\mathbf{Wx})|]$.

The generalization of this idea to **N** dimensions leads to maximizing the expected log determinant of the absolute value of the Jacobian matrix $|[\partial g_i(u_i)/\partial x_j]_{ij}|$. This optimization attempts to map the input vectors uniformly into the unit **N**-cube (assuming that the g-functions are still 0-1 bounded). Intuitively, if the outputs are spread evenly (like molecules of a gas) throughout their (**N**-cube) range, then learning the value of a data point on one axis gives no information about its values on the other axes and maximum independence has been achieved. Bell and Sejnowski [6] showed that the stochastic gradient descent algorithm that maximizes $\mathbf{H}(\mathbf{g}(\mathbf{u}))$ is

$$\Delta \mathbf{W} \propto \mathbf{W}^{-T} + \mathbf{f}(\mathbf{u})\mathbf{x}^T \tag{8}$$

where $-T$ denotes inverse transpose, and the vector-function, $\mathbf{f}$, has elements

$$f_i(u_i) = \frac{\partial}{\partial u_i} \ln g_i'(u_i). \qquad (9)$$

When $g_i'(u_i) = p_i(u_i)$ for all i, then, according to (6), the ICA algorithm is exact. Unfortunately, this leaves a difficulty. Either one has to estimate the functions g during training, or one needs to assume that the final term in (6) does not interfere with Infomax performing ICA. We have empirically observed a systematic robustness to misestimation of the prior, $\hat{p}_i(u_i|\mathbf{W}, g) = |g_i'(u_i)|$. Although unproven, this robustness conjecture can be stated [16]: Any super-Gaussian prior will suffice to extract super-Gaussian independent components. Any sub-Gaussian prior will suffice to extract sub-Gaussian independent components. This conjecture also leads to the generally successful "extended ICA" algorithms [9], [10] that switch the component priors, $\hat{p}_i(u_i)$, between super- and sub-Gaussian functions. In practice, as the robustness principle suggests, this switching may be all the estimation needed to obtain a correct solution. The same insight underlies "negentropy" approaches to ICA that maximize the distance of the $p_i(u_i)$ from Gaussian, described in [13] and by Lee *et al.* [10].

For most natural data (images, sounds, etc.), the independent component pdfs are all super-Gaussian, so many good results have been achieved using "logistic ICA," in which the super-Gaussian prior is the slope, $g_i'(u_i)$, of the common logistic sigmoid function (8) so often used in neural networks. For this choice of g, the function f in (8) evaluates simply to $f(u) = 1 - 2g(u)$.

Infomax ICA is almost identical to the maximum likelihood approach [3]. In maximum likelihood density estimation, one maximizes a parameterized estimate of the log of the pdf of the input, $\log \hat{p}(\mathbf{x}|\mathbf{W}, g)$. A simple argument shows that the absolute value of the determinant of the Jacobian matrix, $|\det[\partial g_i(u_i)/\partial x_j]_{ij}|$ is exactly such a density estimate [for much the same reason that $|g_i'(u_i)|$ is a density estimate for $p_i(u_i)$ in (6)]. Infomax maximizes this log likelihood, and therefore inherits the useful properties of maximum likelihood methods while preserving an information-theoretic perspective on the problem.

An additional and important feature was added to the Infomax ICA algorithm by Amari and colleagues [8], who observed that a simpler learning rule, with much faster and more stable convergence, could be obtained by multiplying the Infomax gradient of (8) by $\mathbf{W}^T\mathbf{W}$, obtaining

$$\Delta \mathbf{W}_{NatGrad} = (\Delta \mathbf{W})\mathbf{W}^T\mathbf{W} \propto (\mathbf{I} + \mathbf{f}(\mathbf{u})\mathbf{u}^T)\mathbf{W}. \quad (10)$$

Since $\mathbf{W}^T\mathbf{W}$, which scales the gradient, is positive-definite, it does not change the minima and maxima of the optimization. Its optimality has been explained using information geometry [8] and equivalence—the gradient vector local to $\mathbf{W}$ is normalized to behave as if it were close to $\mathbf{I}$ (see [14]). Both interpretations reflect the fact that the parameter space of $\mathbf{W}$ is not truly Euclidean, since its axes are entries of a matrix. Equation (10) is clearly a nonlinear

decorrelation rule, stabilizing when $\langle -\mathbf{f}(\mathbf{u})\mathbf{u}^T \rangle = \mathbf{I}$. (The minus sign is required because the $\mathbf{f}$ functions are typically decreasing.) The Taylor series expansion of the $\mathbf{f}$ functions provides information about higher order correlations necessary to perform ICA.

In addition to its effective use in solving blind source separation problems in signal processing where known independent "ground truth" sources are known, at least for test examples, ICA has also been applied to data from the natural world where the degree to which the ICA assumptions are satisfied is unknown and for which no clear idea of what the maximally independent sources may be. First, we will examine an application of ICA to natural images that supports an Infomax-based theory of perceptual brain organization [17] and also illustrates the nature of independence.

III. DECOMPOSITION OF NATURAL IMAGES

The classic experiments of Hubel and Wiesel on neurons in primary visual cortex revealed that many of them are orientation-selective feature detectors. This raised the question "Why do we have edge detectors?" In other words, are there coding principles that predict the formation of localized, oriented receptive fields? Horace Barlow proposed that our visual cortical feature detectors might be the end result of a *redundancy reduction* process, in which the activation of each feature detector is supposed to be as statistically independent from the others as possible. However, algorithms based only on second-order statistics failed to give local filters. In particular, the principal components of natural images are Fourier filters ranked in frequency, quite unlike oriented localized filters. Other researchers have proposed "projection pursuit"-style approaches to this problem, culminating in Olshausen and Field's [18] demonstration of the self-organization of local, oriented receptive fields using a sparseness criterion.

The assumption implicit in this approach is that early visual processing should attempt to invert the simplest possible image formation process, in which the image is formed by linear superposition of basis vectors (columns of $\mathbf{W}^{-1}$), each activated by independent (or sparse) causes, $\mathbf{u}_i$. Bell and Sejnowski [19] showed that ICA basis images for a set of small image patches taken at random from natural images do consist of oriented, localized contrast-sensitive functions ["edge-detectors" (Fig. 3)]. Since sparseness is often related to super-Gaussianity, it is clear why logistic ICA produced filters sensitive to sparse patterns. These distributions, furthest from Gaussian on the super-Gaussian side, are the most likely to be as statistically independent as possible, through the Central Limit Theorem argument that any mixture of two independent distributions should produce a distribution that is closer to Gaussian. Note that none of the independent components of these data were sub-Gaussian, as was verified using the "extended ICA" algorithm [10]. Later, van Hateren and Ruderman derived *basis movies* of moving images (http://hlab.phys.rug.nl/demos/ica) [20], which were localized, oriented and moving perpendicular to their orientation direction, as in monkey visual cortex.
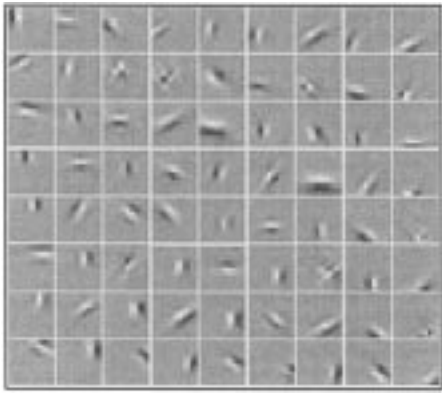
Fig. 3. A selection of 144 basis functions (columns of $\mathbf{W}^{-1}$) obtained from training on patches of 12-by-12 pixels from pictures of natural scenes.

IV. ANALYSIS OF EEG AND AVERAGED EVENT-RELATED POTENTIAL (ERP) DATA

The EEG is a noninvasive measure of brain electrical activity recorded as changes in potential difference between points on the human scalp. Because of volume conduction through cerebrospinal fluid, skull and scalp, EEG data collected from any point on the scalp may include activity from multiple processes occurring within a large brain volume. This has made it difficult to relate EEG measurements to underlying brain processes or to localize the sources of the EEG signals. Furthermore, the general problem of determining the distribution of brain electrical sources from electromagnetic field patterns recorded on the scalp surface is mathematically underdetermined.

Event-related potentials (ERPs), time series of voltages from the ongoing EEG that are time-locked to a set of similar experimental events, are usually averaged prior to analysis to increase their signal/noise relative to other nontime and phase-locked EEG activity and nonneural artifacts. For several decades, ERP researchers have proposed a number of techniques to localize the sources of stimulus-evoked potentials, either by assuming a known or simple spatial configuration [21], or by restricting generator dipoles to lie within and point outward from the cortical surface [22].

Makeig *et al.* [23] noted that ICA could be used to separate the problem of EEG (or MEG) source identification from the problem of source localization. That is, ICA (or other blind source separation algorithm) may tell *what* temporally independent activations compose the collected scalp recordings without specifying directly *where* in the brain these activations arise. By separating the contributions of different brain and nonbrain sources to the data, however, ICA is also proving to be an efficient preprocessing step prior to source localization [24].

Subsequent work has explored the application of ICA to collections of averaged ERPs [11], [25]–[27] to unaveraged single-trial ERP epochs [28]–[32], and to clinical EEG data [33]. Many other groups, including MEG researchers [34]–[37], have now begun working in this area.

A. Assumptions of ICA Applied to EEG/MEG Data

Four main assumptions underlie ICA decomposition of EEG (or MEG) time series. 1) Signal conduction times are equal and summation of currents at the scalp sensors is linear, both reasonable assumptions for currents carried to the scalp electrodes by volume conduction at EEG frequencies [38] or for superposition of magnetic fields at SQUID sensors. 2) Spatial projections of components are fixed across time and conditions. 3) Source activations are temporally independent of one another across the input data. 4) Statistical distributions of the component activation values are not Gaussian.

What are the Independent Components? For biomedical time series analysis (EEG, MEG, etc.), the rows of the input matrix, $\mathbf{x}$, are EEG/ERP signals recorded at different electrodes and the columns are measurements recorded at different time points. ICA finds an "unmixing" matrix, $\mathbf{W}$, that decomposes or linearly unmixes the multichannel scalp data into a sum of temporally independent and spatially fixed components, $\mathbf{u} = \mathbf{Wx}$. The rows of the output data matrix, $\mathbf{u}$, are time courses of activation of the ICA components. The columns of the inverse matrix, $\mathbf{W}^{-1}$, give the relative projection strengths of the respective components at each of the scalp sensors. These scalp weights give the scalp topography of each component, and provide evidence for the components' physiological origins (e.g., eye activity projects mainly to frontal sites). The projection of the ith independent component onto the original data channels is given by the outer product of the ith row of the component activation matrix, $\mathbf{u}$, with the ith column of the inverse unmixing matrix, and is in the original channel locations and units (e.g., μV). Thus, brain activities of interest accounted for by single or by multiple components can be obtained by projecting selected ICA component(s) back onto the scalp, $\mathbf{x_0} = \mathbf{W}^{-1}\mathbf{u_0}$, where $\mathbf{u_0}$ is the matrix, $\mathbf{u}$ of activation waveforms with rows representing activations of irrelevant component activation(s) set to zero.

B. Analyzing Collections of Averaged ERPs

Many studies employ ERP peak measures to test clinical or developmental hypotheses. However, ERPs cannot be easily decomposed into functionally distinct components, because time courses and scalp projections of different brain generators generally overlap. We have shown, however, that ICA can effectively decompose multiple overlapping components from selected sets of related ERP averages [25].

Fig. 4 illustrates results of decomposing a collection of 25 to 75 1-s averages from different task and/or stimulus conditions, each summing a relatively large number of single trials (250–7000). Participating subjects, eight males and two females, were right-handed with normal or corrected to normal vision. During 76-s trial blocks, subjects were instructed to attend to one of five squares continuously displayed on a back background 0.8 cm above a centrally located fixation point. The (1.6 × 1.6 cm) squares were positioned horizontally at angles of $0°$, $\pm 2.7°$, and $\pm 5.5°$ in the visual field $2°$ above from the point of fixation. Four squares were outlined in blue while one, marking the attended location, was outlined in
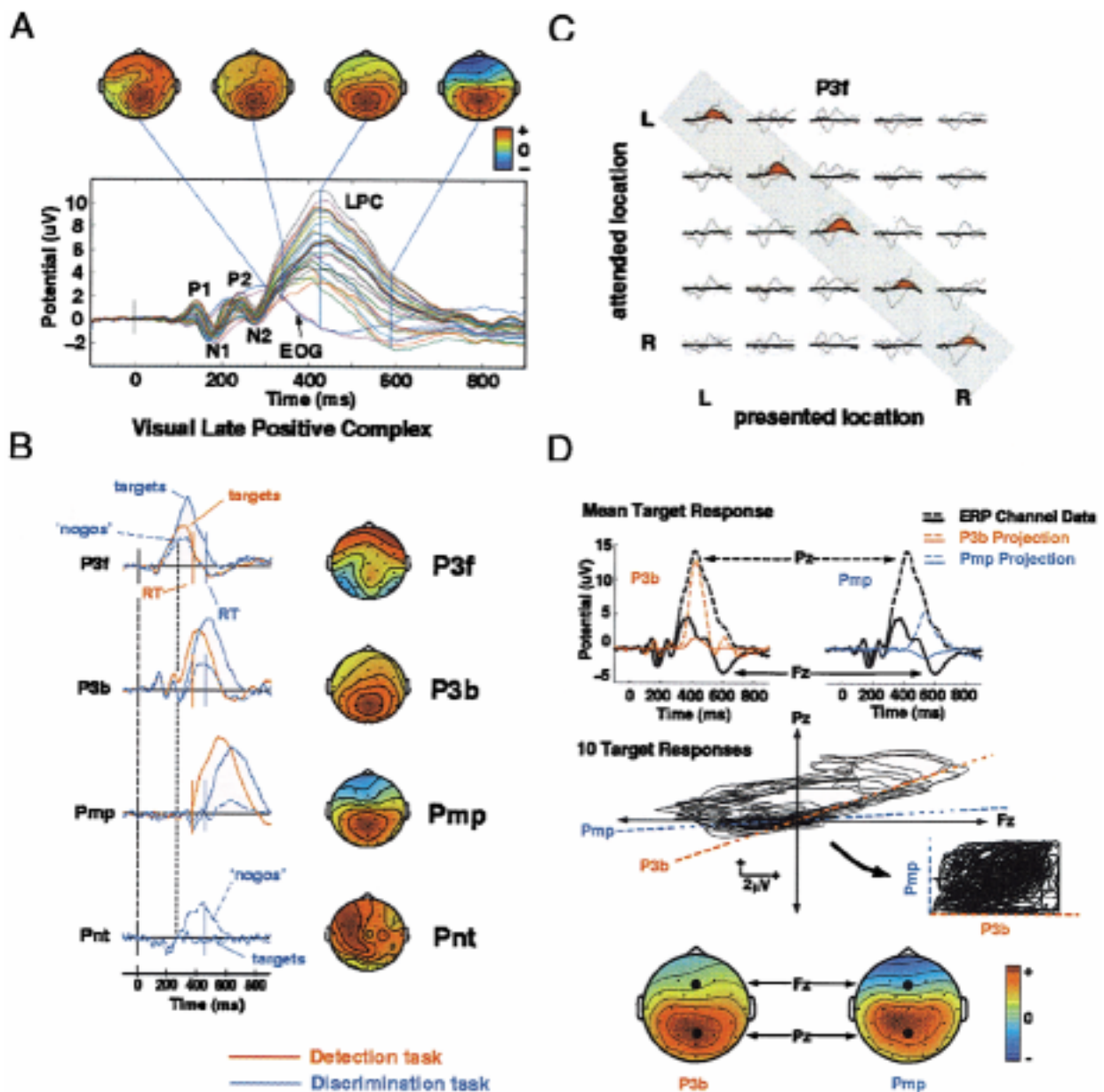
Fig. 4. (a) Grand mean evoked response to detected target stimuli in the detection task (average of responses from ten subjects and five attended locations). Response waveform at all 29 scalp channels and two EOG channels are plotted on a common axis. Topographic plots of the scalp distribution of the response at four indicated latencies show that the LPC topography is labile, presumably reflecting the summation at the electrodes of potentials generated by temporally overlapping activations in several brain areas each having broad but topographically fixed projections to the scalp. All scalp maps shown individually scaled to increase color contrast, with polarities at their maximum projection as indicated in the color bar. (b) Activation time courses and scalp maps of the four LPC components produced by Infomax ICA applied to 75 1-s grand-mean (10-Ss) ERPs from both tasks. Map scaling as in (a). The thick dotted line (*left*) indicates stimulus onset. Mean subject-median response times (RTs) in the Detection task (*red*) and Discrimination task (*blue*) are indicated by solid vertical bars. Three independent components (P3f, P3b, Pmp) accounted for 95%–98% of response variance in both tasks. In both tasks, median RT coincided with Pmp onset. The faint vertical dotted line near 250 ms shows that the P3f time courses for targets and "nogo" nontargets (presented in the target location) just at the onset of the left-sided Pnt component, which was active only in this condition. (c) Envelopes of the scalp projections of maximally independent component P3f, (*red filled*) superimposed on the mean response envelopes (*black outlines*) for all 5 × 5 response conditions of the Detection task. (d) The top panels show the grand mean target response at two scalp channels, Fz and Pz (*thick traces*), and the projections of the two largest ICA components, P3b and Pmp, to the same channels (*thin traces*). The central panel shows a scatter plot of ten average target ERPs at the two electrodes. The data contain two strongly radial (and, therefore, spatially fixed) features. The dashed lines (*middle panel*) show the directions associated with components P3b and Pmp in these data, as determined by the relative projection strengths of each component to these two scalp channels (*shown below as black dots on the component scalp maps*). The degree of data entropy attained by ICA training is illustrated by the (*center right*) plot insert, which shows the (31-channel) scatter-plotted data after logistic transformation and rotation to the two component axes (from [25] by permission).

green. The location of the attended location was counterbalanced across trial blocks. In the "Detection" task condition, all stimuli were filled circles and subjects were required to press a right-hand held thumb button as soon as possible following stimuli presented in the attended location. In the "Discrimination" task condition, 75% of the presented stimuli were filled circles, the other 25% filled squares. Subjects were required to press the response button only in response to filled squares appearing in the attended location, and to ignore filled circles. In this condition, 35 blocks of trials were collected from each subject, seven blocks at each of the five possible attended locations. Each block included 35 target squares and 105 distractor (or "nogo") circles presented at the attended location, plus 560 circles and squares presented at the four unattended locations.

EEG data were collected from 29 scalp electrodes mounted in a standard electrode cap (Electrocap, Inc.) at locations based on a modified International 10–20 System, and from two periocular electrodes placed below the right eye and at the left outer canthus. All channels were referenced to the right mastoid with input impedance less than 5 kΩ. Data were sampled at 512 Hz within an analog passband of 0.01–50 Hz.

Responses evoked by target stimuli [their grand mean shown in Fig. 4(a), *colored traces*] contained a prominent "late positive complex" (LPC, often called "P300") following expected early peaks P1, N1, P2, and N2. In the grand-mean detection-task response, the scalp topography of the response varied continuously [Fig. 4(a), *scalp maps*].

ICA was applied to all 75 31-channel responses from both tasks (1-s ERPs from 25 detection-task and 50 discrimination-task conditions, each a grand average over ten subjects) producing 31 temporally independent components. Of these, just three accounted for 95%–98% of the variance in the ten target responses from both tasks.

Component P3f (blue traces) became active near the N1 peak. Its active period *continued through* the P2 and N2 peaks and the upward slope of the LPC. That is, P3f accounted for a slow shift beginning before LPC onset, positive at periocular and frontal channels and weakly negative at lateral parietal sites (*top rows*). Fig. 4(c) shows the 31-channel projections of P3f as (*red*) filled data envelopes within the outlined envelope of the whole responses in each condition.

Component P3b, the largest of the three independent LPC components, had a central parietal maximum and a right-frontal bias, like the LPC peak itself. In the Detection task, its peak amplitude appeared inversely related to median RT. In the Discrimination task, the ∼90 ms delay between RT and the P3b peak observed in the detection task was reproduced only in the fast-responder response. These characteristics of the central LPC component (P3b) identified by ICA appear consistent with those of the LPC peak in the Detection task. However, in the Discrimination-task subaverages, the ERP and P3b peaks did *not* coincide. The P3b component also accounted for some early response activity. This appeared to reflect a tendency of the algorithm to make very large components "spill over" to account for periods of weak activity with related scalp distributions. Subsequent decompositions of the

Detection-task data by PCA with our without Varimax, and Promax rotation (see [25]) produced P3b analogs in which this "spillover" was stronger than for ICA.

Component Pmp. Pmp was activated only following target stimuli followed by button presses. Its posterior maximum was contralateral to response hand, and its latency and topographic variability across subjects strongly resembled that of the 200-ms post-movement positivity in the voluntary motor response [39]. However, in the discrimination task no Pmp was present in target responses of the five faster responders. Pmp accounted for the later post-response portion of the LPC originally called SW (for "slow wave") [40].

Component Pnt. A fourth LPC component was active only in response to nontarget stimuli presented in the target location. Inhibiting a quick motor response to these stimuli required conscious effort. Component Pnt (for nontarget) had a left-sided scalp projection, consistent with lesion and fMRI studies showing that a brain area below these electrodes is involved in response inhibition.

These results suggested that ICA can parsimoniously decompose complex ERP data sets comprised of many scalp channels, stimulus types, and task conditions into temporally independent, spatially fixed, and physiologically plausible components without necessarily requiring the presence of multiple local response peaks to separate meaningful response components. Although the results reported here and elsewhere [11], [25], [26] are encouraging, we need to keep in mind that for averaged ERP data, the ICA assumptions may only be approximately satisfied. That is, as illustrated in Fig. 4(d), given real data, any ICA algorithm can return at best *maximally* independent components.

Spatial stationarity. Spatial stationarity of the component scalp maps, assumed in ICA, is compatible with the observation made in large numbers of functional imaging reports that performance of particular tasks increases blood flow within small ($\approx$cm^3), discrete brain regions [41]. ERP sources reflecting task-related information processing are generally assumed to sum activity from spatially stationary generators, although stationarity might not apply to subcentimeter scales or to some spontaneous macroscopic EEG phenomena such as spreading depression or sleep spindles [42]. Our results to date suggest that most EEG oscillations, including alpha rhythms, can be better modeled as composed of temporally independent islands of coherent cortical activity, rather than as traveling waves [32].

Temporal independence. ICA assumes that sources of the EEG must be temporally independent. However, brain components of averaged ERPs most often have temporally overlapping active periods. Independence of ERP features may be maximized by, first, sufficiently and systematically varying the experimental stimulus and task conditions, and, next, training the algorithm on the concatenated collection of resulting event-related response averages. Fortunately, the first goal of experimental design, to attain independent control of the relevant output variables, is compatible with the ICA requirement that the activations of the relevant data components be independent. Thus, for example, the subject group-mean ERP data we analyzed successfully using ICA

(Fig. 4) consisted of collections of 25 to 75 1-s averages from different task and/or stimulus conditions, each summing a relatively large number of single trials (250–7000). Unfortunately, however, independent control of temporally overlapping ERP components may be difficult or impossible to achieve. Simply varying stimuli and subject task does not guarantee that all the spatiotemporally overlapping response components appearing in the averaged responses will be activated independently in the resulting data. Thus, the suitability of ICA for decomposition of small sets of ERP averages cannot be assumed, and such decompositions must be examined carefully, using convergent behavioral and/or physiological evidence, before accepting the functional independence of the derived components. Also, ERP components, even those derived by ICA, may actually represent sums of event-related phase and amplitude perturbations in components of the ongoing EEG (see below).

Dependence on source distribution. Because of the tendency described by the Central Limit Theorem, mixtures that appear normally distributed may be the sum of non-Gaussian sources. In theory, ICA cannot separate multiple Gaussian processes, although in practice even small deviations from normality can suffice to give good results. Also, not all ICA algorithms are capable of unmixing independent components with sub-Gaussian (negative-kurtosis) distributions. For example, the original Infomax ICA algorithm using the logistic nonlinearity is biased toward finding super-Gaussian (sparsely activated) independent components (i.e., sources with positive kurtosis). Super-Gaussian sources, which have more near-zero values than the best-fitting Gaussian process, are common in speech and many other natural sounds and visual images (see Section III) [19], [43].

The assumption that source distributions are super-Gaussian is compatible with the physiologically plausible assumption that an averaged ERP is composed of one or more overlapping series of relatively brief activations within spatially fixed brain areas performing separable stages of stimulus information processing. Nonetheless, sub-Gaussian independent components have been demonstrated in EEG data [28], including line noise, sensor noise and low frequency activity. In practice, however, sub-Gaussian components rarely appear in ERPs or in spontaneous EEG. The super-Gaussian statistics of independent components of ERP data may indicate that brain information processing is dominated by spatially sparse, intermittently synchronous brain structures.

C. Analyzing Collections of Event-Related EEG Epochs

Response averaging ignores the fact that response activity may vary widely between trials in both time course and scalp distribution. This temporal and spatial variability may in fact reflect changes in subject performance or in subject state (possibly linked to changes in attention, arousal, task strategy, or other factors). Thus conventional averaging methods may not be suitable for investigating brain dynamics arising from intermittent changes in subject state and/or from complex interactions between task events. Further, response averaging makes possibly unwarranted assumptions about the relationship between ERP features and the dynamics of the ongoing EEG.

Analysis of single event-related trial epochs may potentially reveal more information about event-related brain dynamics than simple response averaging, but faces three signal processing challenges: 1) difficulties in identifying and removing artifacts associated with blinks, eye-movements and muscle noise, which are a serious problem for EEG interpretation and analysis; 2) poor signal-to-noise ratio arising from the fact that nonphase locked background EEG activities often are larger than phase-locked response components; and 3) trial-to-trial variability in latencies and amplitudes of both event-related responses and endogenous EEG components. Additional interest in analysis of single-trial event-related EEG (or MEG) epochs comes from the realization that filtering out time- and phase-locked activity (by response averaging) isolates only a small subset of the actual event-related brain dynamics of the EEG signals themselves [44].

Recently, a set of promising analysis and visualization methods for multichannel single-trial EEG records have been developed that may overcome these problems [23], [29], [45]. These tools were first used to analyze data from the aforementioned visual Detection experiment on 28 control subjects plus 22 neurological patients whose EEG data, recorded at 29 scalp and two EOG sites, were often heavily contaminated with blink and other eye-movement artifacts.

To visualize collections of single-trial EEG records, "ERP image" plots [29], [45] are useful and often reveal unexpected inter-trial consistencies and variations. Fig. 5(a) shows all 641 single-trial ERP epochs recorded from an autistic subject time-locked to onsets of target stimuli (*left vertical line*). Single-trial event-related EEG epochs recorded at the vertex (Cz) and at a central parietal (Pz) site are plotted as color-coded horizontal traces (see color bar) sorted in order of the subject's reaction time latencies (*thick black line*). The ERP average of these trials is plotted below the ERP image.

ICA, applied to these 641 31-channel EEG records, separated out (*clockwise*): 1) artifact components arising from blinks or eye movements, whose contributions could be removed from the EEG records by subtracting the component projection from the data [30], [46]; 2) components showing stimulus time-locked potential fluctuations of consistent polarity many or all trials; 3) components showing response-locked activity covarying in latency with subject response times; 4) "mu-rhythm" components [47] at approximately 10 Hz that decreased in amplitude when the subject responds; 5) other components having prominent alpha band (8–12 Hz) activity whose intertrial coherence [Fig. 5(b), *lower middle panel, bottom trace*], measuring phase-locking to stimulus onsets, increased significantly after stimulus presentation, even in the absence of any alpha band power increase (*middle trace*); and 6) other
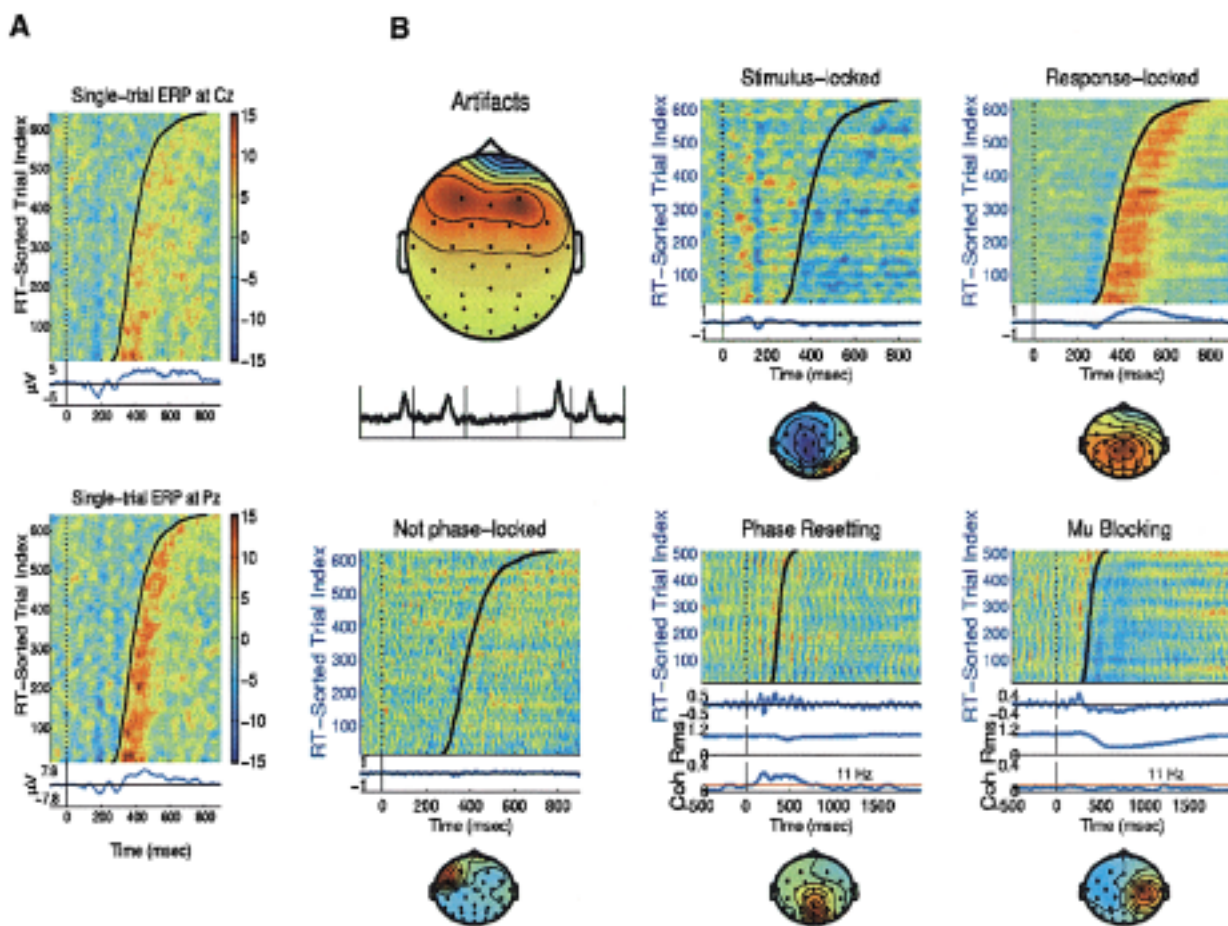
Fig. 5. ERP-image plots of target response data from a visual selective attention experiment and various independent component categories. (a) Single-trial ERPs recorded at a central (Cz) and a parietal electrode (Pz) from an autistic subject and timelocked to onsets of visual target stimuli (*left thin vertical line*) with superimposed subject response times (RT). (b) Single-trial activations of sample independent components accounting for (*clockwise*) eye blink artifacts, stimulus-locked and response-locked ERP components, response-blocked oscillatory mu, stimulus phase-reset alpha, and nonphase locked activities.

EEG components whose activities were either unaffected by experimental events or were affected in ways not revealed by these measures. This taxonomy could not have been obtained from signal averaging or other conventional frequency-domain approaches.

Better understanding of trial-to-trial changes in brain responses may allow a better understanding of normal human performance in repetitive tasks, and a more detailed study of changes in cognitive dynamics in normal, brain-damaged, diseased, aged, or genetically abnormal individuals. ICA-based analysis also allows investigation of the interaction between phenomena seen in ERP records and its origins in the ongoing EEG. Contrary to the common supposition that ERPs are brief stereotyped responses elicited by some events and independent of ongoing background EEG activity, many ERP features may be generated by ongoing EEG processes (see Fig. 5).

Decomposition of unaveraged single-trial EEG records allows: 1) removal of pervasive artifacts from single-trial EEG records, making possible analysis of highly contaminated EEG records from clinical populations [46]; 2) identification and segregation of stimulus- and response-locked

EEG components; 3) realignment of the time courses of response-locked components to prevent temporal smearing in the average; 4) investigation of temporal and spatial variability between trials; and 5) separation of spatially overlapping EEG activities that may show a variety of distinct relationships to task events. The ICA-based analysis and visualization tools increase the amount and quality of information in event- or response-related brain signals that can be extracted from event-related EEG (or MEG) data. ICA thus may help researchers to take fuller advantage of what until now has been an only partially realized strength of event-related paradigms—the ability to examine systematic relationships between single trials within subjects [29], [32], [45], [48].

Although ICA appears to be a promising method for analyzing for EEG and MEG data, results of ICA must be interpreted with caution. In general, the effective number of independent components contributing to the scalp EEG is unknown and between-subject variability is difficult to resolve. One approach is to sort components into between-subject clusters recognizable by their spatial and temporal patterns as well as by their time-domain (ERP) and frequency-domain
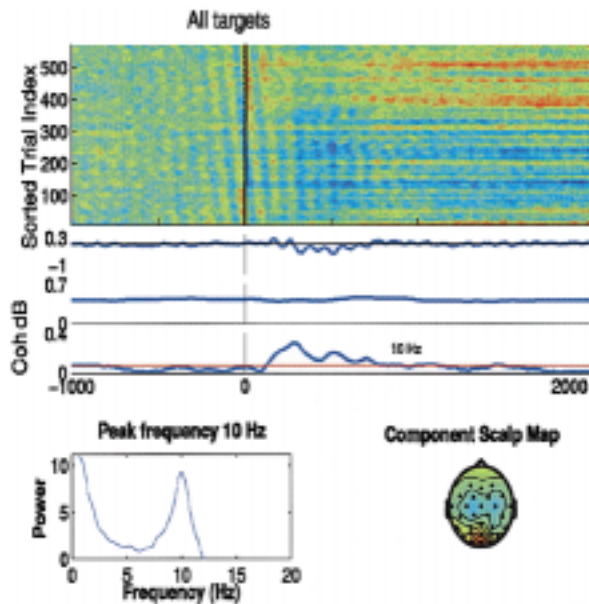
Fig. 6. ERP-image plot of single-trial activations of one alpha component from the selective visual attention experiment described in Section IV. *Top image*: Single-trial potentials, color coded (*scale*: *red* positive, *green* zero and *blue* negative). Blue traces below image: (*top trace*) averaged evoked response activity of this component, showing "alpha ringing." *Units*: proportional to μV. (*middle trace*) Time course of rms amplitude of this component at its peak frequency, 10 Hz. *Units*: relative to $\log_{10}(\mu V^2)$. (*bottom trace*) Time course of inter-trial coherence at 10 Hz. (*thick*), plus the bootstrap ($p = 0.02$) significance threshold (*thin*). Intertrial coherence measures the tendency for phase values at a given time and frequency to be fixed across trials. *Bottom left*: Mean power spectral density of the component activity (units, relative decibels). *Bottom right*: scalp map showing the interpolated projection of the component to the scalp electrodes.

(e.g., ERSP, event-related spectral perturbation) reactivities [32].

D. Case Study: Stimulus-Induced "Alpha Ringing"

EEG data were recorded from a subject performing the selective attention experiment described above. Fig. 6 shows an ERP-image plot giving the time course of activation of one independent component whose activity spectrum (*lower left*) had a strong peak (near 10 Hz) in the alpha range. Its map (*lower right*) could be well approximated by the projection of a single equivalent dipole, suggesting that its source might be a small patch of unknown size in left medial occipital cortex.

The "ERP-image" view shows the time course of activation of this component in over 500 single trials each time locked to the presentation of a target stimulus (*vertical line*). Here, the trials have been sorted not in order of response time (as in Fig. 5), but rather in order of their phase at 10 Hz in a three-cycle window ending at stimulus onset. Phase sorting produces an apparent autocorrelation of the signals. The ERP (*uppermost trace*) suggests that following the stimulus this component produced nearly a second of increased alpha activity superimposed on a slow negative wave. Note, however, that the slope of the negative-phase lines (*dark stripes*) becomes near-vertical a half second (*at the first tick mark*) after stimulus presentation. This change in slope represents a sys-

tematic *phase resetting* of the component alpha activity induced by the stimulus.

The vertically time-aligned phase maxima between 200 and 700 ms after the stimulus produces the appearance of increased 10-Hz activity in the ERP (*upper trace*). However (as the *middle trace* shows), mean power at 10 Hz in the single-trial EEG itself does *not* increase above its baseline. Instead (as the *lower trace* shows), the phase resetting of the component process by the stimulus, becomes significant (*horizontal thin line*) about 200 ms after stimulus onset, producing the 10-Hz ripples in the ERP.

Here, ICA allows the actual event-related EEG dynamics producing the observed (normal) "alpha-ringing" in the averaged evoked response to be accurately modeled, whereas measuring the average evoked response alone could suggest a quite different (and incorrect) interpretation. As Makeig *et al.* [32] have shown, ICA identifies several clusters of independent EEG alpha components, each with a scalp map resembling the projection of one (or in one cluster of cases, two) dipoles located in the posterior cortex. Typically, several of these sum to form a subject's spatially complex and variable recorded "alpha rhythm."

E. Component Stability

We have investigated the component stability of ICA decomposition of EEG/ERPs at three different scales: 1) the replicability of components from repeated ICA trainings on the same data set; 2) within-subject spatiotemporal stability of independent components of collections of temporally overlapping or distinct subepochs of the single-trial EEG records; and 3) between-subject replicability of independent components of 1- or 3-s single-trial EEG epochs.

Infomax ICA decomposition is relatively robust and insensitive to the exact choice or learning rate or data batch size [25], [26], [31], [49]. Training with data in different random orders has little effect on the outcome: Independent components with large projections are stable, though typically the smallest components vary.

The within-subject spatiotemporal stability of independent components of the EEG was tested by applying moving-window ICA to overlapping event-related subepochs of the same single-trial recordings used in this study [32]. ICA decomposition of multichannel single-trial event-related EEG data gave stable and reproducible spatiotemporal structure and dynamics of the EEG before, during and after experimental events. In addition, component clusters identified in single ICA decompositions of concatenated whole target-response epochs strongly resembled those produced by the subwindow decompositions.

We then investigated the between-subject stability of independent components of the single-trial EEG epochs by applying a component clustering analysis (component-matching based on the component scalp maps and power spectra of component activations) to 713 components derived from 23 normal controls participating in the same visual selective attention task. Clusters accounting for eye blinks, lateral eye movement, and temporal muscle activities

contained components from almost all subjects. In general, clusters accounting for early stimulus-locked activity, late response-locked activity (P300), event-modulated mu and alpha band activities (see Fig. 5) were largely replicated in many subjects.

V. FUNCTIONAL MAGNETIC RESONANCE IMAGE

The fMRI technique is a noninvasive technique making it possible to localize dynamic brain processes in intact living brains [50]. It is based on the magnetic susceptibilities of oxygenated hemoglobin (HbO_2) and deoxygenated hemoglobin (HbR) and is used to track blood-flow-related phenomena accompanying or following neuronal activations. The most commonly used fMRI signal is the blood-oxygen-level-dependent (BOLD) contrast [51]. The analysis of fMRI brain data is a challenging enterprise, as the fMRI signals have varied, unpredictable time courses that represent the summation of signals from hemodynamic changes as a result of neural activities, from subject motion and machine artifacts, and from physiological cardiac, respiratory and other pulsations. The relative contribution and exact form of each of these components in a given session is largely unknown to the experimenter, suggesting a role for blind separation methods if the data have properties that are consistent with these models [52]–[55].

The assumptions of ICA apply to fMRI data in a different way than to other time series analysis. Here, the principle of brain modularity suggests that, as different brain regions perform distinct functions, their time courses of activity should be separable (though not necessarily independent, particularly when, typically, only a few hundred or fewer time points are available). Spatial modularity, plus the relatively high 3-D spatial resolution of fMRI, allows the use of ICA to identify maximally *spatially* independent regions with distinguishable time courses. Decreases as well as increases in brain activity are observed, which allows components to have overlapping spatial regions and still be approximately independent. However, the spatial independence of active brain areas is not perfect, and therefore the nature and functional significance of independent fMRI components must be validated by convergent physiological and behavioral evidence.

A. General Linear Model (GLM)

Traditional methods of fMRI analysis [41] are based on variants of the general linear model (GLM), i.e.,

$$\mathbf{X} = \mathbf{G}\beta + \varepsilon \tag{11}$$

where

X	n by v row mean-zero data matrix with n the number of time points in the experiment;
v	total number of voxels in all slices;
$\mathbf{G}$	specified n by p design matrix containing the time courses of all p factors hypothesized to modulate the BOLD signal, including the behavioral manipulations of the fMRI experiment;
β	p by v matrix of parameters to be estimated;
ε	matrix of noise or residual errors typically assumed to be independent, zero-mean and Gaussian distributed, i.e., $\mathbf{N}(0, \sigma^2)$.

Once $\mathbf{G}$ is specified, standard regression techniques can be used to provide a least squares estimate for the parameters in β. The statistical significance of these parameters can be considered to constitute spatial maps [41], one for each row in β, which correspond to the time courses specified in the columns of the design matrix. The GLM assumes: 1) the design matrix is known exactly; 2) time courses are white; 3) the βs follow a Gaussian distribution; and 4) the residuals are thus well-modeled by Gaussian noise.

B. ICA Applied to fMRI Data

Using ICA, we can calculate an unmixing matrix, $\mathbf{W}$, to calculate spatially independent components

$$\mathbf{C} = \mathbf{W}\mathbf{X} \tag{12}$$

where again, $\mathbf{X}$ is the n by v row mean-zero data matrix with n the number of time points in the experiment and v the total number of voxels. $\mathbf{W}$ is an n by n unmixing matrix, and $\mathbf{C}$ is an n by v matrix of n spatially independent component maps (sICs).

If $\mathbf{W}$ is invertible, we could write

$$\mathbf{X} = \mathbf{W}^{-1}\mathbf{C}. \tag{13}$$

An attractive interpretation of (13) is that the columns of $\mathbf{W}^{-1}$ represent temporal basis waveforms used to construct the observed voxel time courses described in the columns of $\mathbf{X}$. Since the rows of $\mathbf{C}$ are maximally independent, the spatial projection of any basis waveform in the data is maximally independent of the spatial projection of any other.

The similarity between ICA and the GLM can be seen by comparing (11) and (13). Starting with (13) and performing the initial simple notation substitutions, $\mathbf{W}^{-1} \rightarrow \mathbf{G}$ and $\mathbf{C} \rightarrow \beta$, we have

$$\mathbf{X} = \mathbf{G}\beta \tag{14}$$

which is equivalent to (11) without the Gaussian error term. Note, however, the important teleological differences between (11) and (14): When regression equation (11) is used, the design matrix $\mathbf{G}$ is specified by the examiner, while in (14) the matrix $\mathbf{G}$, computed from the data by the ICA algorithm, also determines β. That is, ICA does not reply on the experimenter's *a priori* knowledge or assumptions about the time courses of brain activities and recording noise during the recording, and makes only weak assumptions about their probability distributions.

C. An fMRI Case Study

Fig. 7 shows the results of applying ICA to an fMRI data set. The fMRI data were acquired while a subject performed 15-s blocks of visually cued or self-paced right wrist supinations/pronations, alternating with 15-s blocks in which the subject rested. ICA found a maximally spatially independent component that was active during both modes of motor activity and inactive during rest [Fig. 7(b)]. Fig. 7(c) shows re-
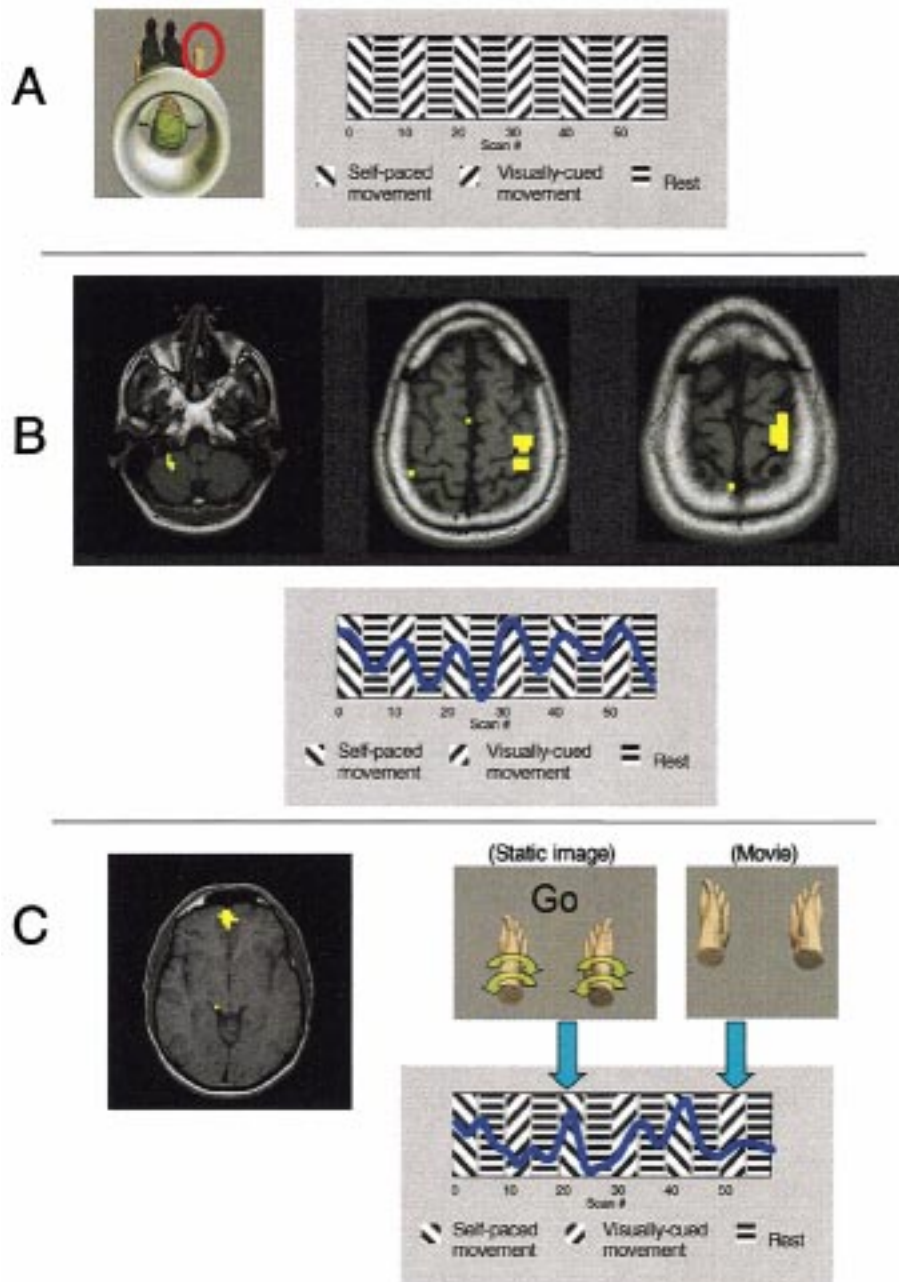
Fig. 7. (a) An fMRI experiment was performed in which the subject was instructed to perform 15-s blocks of alternating right wrist supination and pronation, alternating with 15-s rest blocks. The movement periods where alternately self-paced or were visually cued by a movie showing a moving hand. (b) ICA analysis of the experiment detected a spatially independent component that was active during both types of motor periods but not during rest. The spatial distribution of this component (threshold, $z \geq 2$) was in the contralateral primary motor area and ipsilateral cerebellum. (The radiographic convention is used here, the right side of the image corresponding to the left side of the brain and vice versa) (from McKeown, *et al.*, manuscript in preparation). (c) A similar fMRI experiment was performed in which the subject supinated/pronated both wrists simultaneously. Here, ICA detected a component that was more active during self-paced movements than during either visually cued movement or rest periods. The midline region depicted threshold, $z \geq 2$ is consistent with animal studies showing relative activation of homologous areas during self-paced but not visually cued tasks. (e.g. [69]).

sults from a similar fMRI experiment in which the subject was asked to supinate/pronate both wrists simultaneously. Here, ICA detected a component more active during self-paced movements than during either visually cued movement blocks or rest periods. Its midline, frontal polar location (depicted) is consistent with animal studies showing relative ac-

tivation in this area during self-paced but not during visually cued tasks.

D. Future Directions

In many respects, uses for the GLM and ICA are complementary [56], [57]. The advantage of the GLM is that

it allows the experimenter (given several statistical assumptions) to check the statistical significance of activation corresponding to the experimental hypothesis. The disadvantages of the GLM are related to the fact that these assumptions outlined do not fairly represent the fMRI data. Also, dynamic, distributed patterns of brain activity [58] may not be well modeled by the GLM regression framework, which incorrectly considers each voxel to be a discrete, independent unit.

ICA, on the other hand, has proved to be an effective method for detecting unknown unanticipated activations [52]–[54], [59]–[61] without requiring *a priori* assumptions of time courses and spatial distributions of different brain processes, but does not provide a significance estimate for each activation, which makes difficult for experimenters to interpret their results. McKeown has recently proposed a method that uses ICA to characterize portions of the data and then enables the experimenter to test hypotheses in the context of this data-defined characterization [55] by defining a metric that allows a qualitative assessment of the relative mismatch between hypothesis and data.

VI. DISCUSSION

Biomedical signals are a rich source of information about physiological processes, but they are often contaminated with artifacts or noise and are typically mixtures of unknown combinations of sources summing differently at each of the sensors. Further, for many data sets even the nature of the sources is an open question. Here, we have focused on applications of ICA to analyze EEG and fMRI signals. ICA has also been applied to MEG recordings [37] which carry signals from brain sources and are in part complementary to EEG signals, and to data from positron emission tomography (PET), a method for following changes in blood flow in the brain on slower time scales following the injection of radioactive isotopes into the bloodstream [62]. Other interesting applications of ICA are to the electrocorticogram (EcoG)—direct measurements of electrical activity from the surface of the cortex [63], and to optical recordings of electrical activity from the surface of the cortex using voltage-sensitive dyes [64]. Finally, ICA has proven effective at analyzing single-unit activity from the cerebral cortex [65], [66] and in separating neurons in optical recordings from invertebrate ganglia [67]. Early clinical research applications of ICA include the analysis of EEG recordings during epileptic seizures [33].

In addition to the brain signals that were the focus of this paper, signals from others organs, including the heart [31] and endocrine system [68] have similar problems with artifacts that could also benefit from ICA. ICA holds great promise for blindly separating artifacts from relevant signals and for further decomposing the mixed signals into subcomponents that may reflect the activity of functionally distinct generators of physiological activity.

Strength of ICA Applied to EEG/ERP and fMRI Data

ICA of single-trial or averaged ERP data allows blind separation of multichannel complex EEG data into a sum of temporally independent and spatially fixed components. Our results show that ICA can separate artifactual, stimulus-locked, response-locked, and nonevent related background EEG activities into separate components, allowing:

1) removal of pervasive artifacts of all types from single-trial EEG records, making possible analysis of highly contaminated EEG records from clinical populations;
2) identification and segregation of stimulus- and response-locked event-related activity in single-trail EEG epochs;
3) separation of spatially overlapping EEG activities over the entire scalp and frequency band that may show a variety of distinct relationships to task events, rather than focusing on activity at single frequencies in single scalp channels or channel pairs;
4) investigation of the interaction between ERPs and ongoing EEG.

Historically, there has been a separation between related research in ERPs and EEG. The ERP community has largely ignored the interaction between ERPs and ongoing EEG, whereas the EEG community has primarily analyzed EEG signals in the frequency domain, most often using measures of power in standardized frequency bands. Researchers studying averaged ERP waveforms assume that evoked responses are produced by brief synchronous neural activations in brain areas briefly engaged in successive stages of stimulus-related information processing. In this view, response averaging removes background EEG activity since its phase distribution is independent of experimental events. Our recent results [32], on the contrary, showed that many features of an evoked response may actually be produced by event-related changes in the autocorrelation and cross-correlation structure of ongoing EEG processes, each reflecting synchronous activity occurring continuously in one or more brain regions, or by more subtle perturbations in their dynamics. These new insights would have been difficult to obtain without first separating spatially overlapping stimulus-locked, response-locked event-related activity and event-modulated oscillatory activity into different components by ICA in single-trial EEG epochs. Applying these new techniques reveals that the EEG (and MEG) data are a rich source of information about mechanisms of neural synchronization within and between brain areas.

ICA, applied to fMRI data, has proven to be a powerful method for detecting task-related activations, including unanticipated activations [52]–[54], [59]–[61] that could not be detected by standard hypothesis-driven analyses. This may expand the types of fMRI experiments that can be performed and meaningfully interpreted.

Limitations of ICA Applied to EG and fMRI Data

Although ICA appears to be generally useful for EEG and fMRI analysis, it also has some inherent limitations.

First, ICA can decompose at most N sources from data collected at N scalp electrodes. Usually, the effective number of statistically independent signals contributing to the scalp EEG is unknown, and it is likely that observed brain activity

arises from more physically separable effective sources than the available number of EEG electrodes. To explore the effects of a larger number of sources on the results of the ICA decomposition of a limited number of available channels, we have analyzed simulated EEG recordings generated from a head model and dipole sources that include intrinsic noise and sensor noise [63]. This identifies the conditions when ICA fails to separate correlated sources of ERP signals. Results confirmed that the ICA algorithm can accurately identify the time courses of activation and the scalp topographies of relatively large and temporally independent sources from simulated scalp recordings, even in the presence of a large number of simulated low-level source activities. Another approach to validating ICA is to simultaneously record and compare more than one type of signal, such as concurrent EEG and fMRI which, respectively, have good spatial (fMRI) and temporal resolution (EEG) [70], if these prove to be correlated. However, very small or briefly active EEG sources may be too numerous to separate, particularly in data sets with large numbers of electrodes in which small artifacts may be abundant.

Second, the assumption of temporal independence used by ICA cannot be satisfied when the training data set is small, or when separate topographically distinguishable phenomena nearly always co-occur in the data. In the latter case, simulations show that ICA may derive single components accounting for the co-occurring phenomena, along with additional components accounting for their brief periods of separate activation [63]. Such confounds imply that behavioral or other experimental evidence must be obtained before concluding that ICA components with spatiotemporally overlapping projections are functionally distinct. Independent components may be considered functionally distinct when they exhibit distinct reactivities to experimental events, or when their activations correspond to otherwise observable signal sources.

Third, ICA assumes that physical sources of artifacts and cerebral activity are spatially fixed over time. In general, there is no reason to believe that cerebral and artifactual sources in the spontaneous EEG might not move over time. However, in our data, the relatively small numbers of components in the stimulus-locked, response-locked, and nonphase locked categories, each accounting for activity occurring across sets of 500 or more 1-s trials, suggests that the ERP features of our data were primarily stationary, consistent with repeated observations in functional brain imaging experiments that discrete and spatially restricted areas of cortex are activated during task performance [71]. One general observation that has emerged from applying ICA to brain data is the effectiveness of the independence assumption. In the case of ERP and EEG analysis, the largest components had scalp maps that could be accurately fitted with one or two dipoles. This is unlikely to occur unless time courses of coherently synchronous neural activity in patches of cortical neuropil generating the EEG are nearly independent of one another [32].

Fourth, applied to fMRI data, ICA does not provide an experimenter with a significance estimate for each compo-

nent activation, which may discourage experimenters from attempting to interpret the results. By placing ICA in a regression framework, it is possible to combine some of the benefits of ICA with the hypothesis-testing approach of the GLM [55].

Although results of applying ICA to biomedical signals have already shown great promise and given new insights into brain function, the analysis of these results is still in its infancy. They must be validated using other direct or convergent evidence (such as behavior and/or other physiological measurements) before we can interpret their functional significance. Current research on ICA algorithms is focused on incorporating domain-specific constraints into the ICA framework. This would allow information maximization to be applied to the precise form and statistics of biomedical data.

REFERENCES

[1] J. Herault and C. Jutten, "Space or time adaptive signal processing by neural network models," presented at the Neural Networks for Computing: AIP Conf., 1986.
[2] C. Jutten and J. Herault, "Blind separation of sources I. An adaptive algorithm based on neuromimetic architecture," *Signal Process.*, vol. 24, pp. 1–10, 1991.
[3] D. T. Pham, P. Garat, and C. Jutten, "Separation of a mixture of independent sources through a maximum likelihood approach," presented at the Proc. EUSIPCO, 1992.
[4] P. Comon, "Independent component analysis, a new concept?," *Signal Process.*, vol. 36, pp. 287–314, 1994.
[5] A. Cichocki, R. Unbehauen, and E. Rummert, "Robust learning algorithm for blind separation of signals," *Electron. Lett.*, vol. 30, pp. 1386–1387, 1994.
[6] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Comput.*, vol. 7, pp. 1129–1159, 1995.
[7] J. F. Cardoso and B. H. Laheld, "Equivariant adaptive source separation," *IEEE Trans. Signal Processing*, vol. 44, pp. 3017–3030, 1996.
[8] S. Amari, "Natural gradient works efficiently in learning," *Neural Comput.*, vol. 10, pp. 251–276, 1998.
[9] M. Girolami, "An alternative perspective on adaptive independent component analysis algorithm," *Neural Comput.*, vol. 10, pp. 2103–2114, 1998.
[10] T. W. Lee, M. Girolami, and T. J. Sejnowski, "Independent component analysis using an extended infomax algorithm for mixed sub-Gaussian and superGaussian sources," *Neural Comput.*, vol. 11, pp. 417–441, 1999.
[11] 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," *Proc. Nat. Acad. Sci. USA*, vol. 94, pp. 10 979–10 984, 1997.
[12] T. W. Lee, M. Girolami, A. J. Bell, and T. J. Sejnowski, "A unifying information-theoretic framework for independent component analysis," *Comput. Math. Applicat.*, vol. 39, pp. 1–21, 2000.
[13] A. Hyvaerinen, "Survey on independent component analysis," *Neural Comput. Sur.*, vol. 2, pp. 94–128, 1999.
[14] S. Haykin, *Unsupervised Adaptive Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 2000.
[15] M. Girolami, *Advances in Independent Component Analysis*. New York: Springer-Verlag, 2000.
[16] A. J. Bell, "Independent component analysis," in Handbook of Neural Networks, M. Arbib, Ed., to be published.
[17] A. J. Bell and T. J. Sejnowski, "The 'independent components' of natural scenes are edge filters," *Vision Res.*, vol. 37, pp. 3327–3338, 1997.
[18] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, pp. 607–609, 1996.
[19] A. J. Bell and T. J. Sejnowski, "Edges are the independent components of natural scenes," *Adv. Neural Inform. Process. Syst.*, vol. 9, pp. 831–837, 1997.

[20] J. H. van Hateren and D. L. Ruderman, "Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex," *Proc. R. Soc. London B, Biol. Sci.*, ser. , vol. 265, pp. 2315–2320, 1998.

[21] M. Scherg and D. Von Cramon, "Evoked dipole source potentials of the human auditory cortex," *Electroencephalogr. Clin. Neurophysiol.*, vol. 65, pp. 344–360, 1986.

[22] A. Liu, J. Belliveau, and A. Dale, "Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations," in *Proc. Nat. Academy of Sciences USA*, vol. 95, 1998, pp. 8945–8950.

[23] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," *Adv. Neural Inform. Process. Syst.*, vol. 8, pp. 145–151, 1996.

[24] L. Zhukov, D. Weinstein, and C. Johnson, "Independent component analysis for EEG source localization," *IEEE Eng. Med. Biol. Mag.*, vol. 19, pp. 87–96, 2000.

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

[26] S. Makeig, M. Westerfield, J. Townsend, T.-P. Jung, E. Courchesne, and T. J. Sejnowski, "Functionally independent components of early event-related potentials in a visual spatial attention task," *Philos. Trans. R. Soc. London B, Biol. Sci.*, vol. 354, pp. 1135–1144, 1999.

[27] T.-P. Jung, C. Humphries, T.-W. Lee, S. Makeig, M. J. McKeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts: Comparison between ICA and PCA," *Neural Netw. Signal Process.*, vol. VIII, pp. 63–72, 1998.

[28] ——, "Extended ICA removes artifacts from electroencephalographic data," *Adv. Neural Inform. Process. Syst.*, vol. 10, pp. 894–900, 1998.

[29] T.-P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Analyzing and visualizing single-trial event-related potentials," *Adv. Neural Inform. Process. Syst.*, vol. 11, pp. 118–124, 1999.

[30] T.-P. Jung, S. Makeig, C. Humphries, T. W. Lee, M. J. McKeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, pp. 163–178, 2000.

[31] T.-P. Jung, S. Makeig, T.-W. Lee, M. J. McKeown, G. Brown, A. J. Bell, and T. J. Sejnowski, "Independent component analysis of biomedical signals," in *2nd Int. Workshop on Independent Component Analysis and Signal Separation*, 2000, pp. 633–644.

[32] S. Makeig, S. Enghoff, T.-P. Jung, and T. J. Sejnowski, "Moving-window ICA decomposition of EEG data reveals event-related changes in oscillatory brain activity," in *2nd Int. Workshop on Independent Component Analysis and Signal Separation*, 2000, pp. 627–632.

[33] M. J. McKeown, C. Humphries, V. Iragui, and T. J. Sejnowski, "Spatially fixed patterns account for the spike and wave features in absence seizures," *Br. Topogr.*, vol. V12, pp. 107–116, 1999.

[34] A. K. Barros, R. Vigário, V. Jousmäki, and N. Ohnishi, "Extraction of event-related signals from multichannel bioelectrical measurements," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 583–588, 2000.

[35] S. Ikeda and K. Toyama, "Independent component analysis for noisy data—MEG data analysis," *Neural Networks*, vol. 13, pp. 1063–1074, 2000.

[36] A. C. Tang, B. A. Pearlmutter, and M. Zibulevsky, "Blind source separation of multichannel neuromagnetic responses," *Comput. Neurosci. Proc. Neurocomput.*, vol. 32–33, pp. 1115–1120, 2000.

[37] R. Vigário, J. Särelä, V. Jousmäki, M. Hämäläinen, and E. Oja, "Independent component approach to the analysis of EEG and MEG recordings," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 589–593, 2000.

[38] P. Nunez, *Electric Fields of the Brain*. London, U.K.: Oxford Univ. Press, 1981.

[39] S. Makeig, M. Mueller, and B. Rockstroh, "Effects of voluntary movements on early auditory brain responses," *Exp. Br. Res.*, vol. 110, pp. 487–492, 1996.

[40] R. Simson, H. G. Vaughn Jr., and W. Ritter, "The scalp topography of potentials in auditory and visual discrimination tasks," *Electroencephalogr. Clin. Neurophysiol.*, vol. 42, pp. 528–535, 1977.

[41] K. J. Friston, "Statistical parametric mapping and other analyzes of functional imaging data," in *Brain Mapping, The Methods*, A. W. Toga and J. C. Mazziotta, Eds. San Diego: Academic, 1996, pp. 363–396.

[42] M. J. McKeown, C. Humphries, P. Achermann, A. A. Borbely, and T. J. Sejnowski, "A new method for detecting state changes in the EEG: Exploratory application to sleep data," *J. Sleep Res.*, vol. V7, pp. 48–56, 1998.

[43] A. J. Bell and T. J. Sejnowski, "Learning the higher-order structure of a natural sound," *Netw. Comput. Neural Syst.*, vol. 7, pp. 261–266, 1996.

[44] S. Makeig, "Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones," *Electroencephalogr. Clin. Neurophysiol.*, vol. 86, pp. 283–293, 1993.

[45] T.-P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Analysis and visualization of single-trial event-related potentials," *Hum. Br. Map.*, to be published.

[46] 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," *Clin. Neurophysiol.*, vol. 111, pp. 1745–1758, 2000.

[47] S. Makeig, S. Enghoff, T.-P. Jung, and T. J. Sejnowski, "A natural basis for brain-actuated control," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 208–211, 2000.

[48] K. Kobayashi, C. J. James, T. Nakahori, T. Akiyama, and J. Gotman, "Isolation of epileptiform discharges from unaveraged EEG by independent component analysis," *Clin. Neurophysiol.*, vol. 110, pp. 1755–1763, 1999.

[49] S. Makeig, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of single-trial event-related potentials," *Soc. Neurosci. Abstr.*, vol. 23, p. 1061, 1997.

[50] K. K. Kwong, J. W. Belliveau, D. A. Chesler, I. E. Goldberg, R. M. Weisskoff, B. P. Poncelet, D. N. Kennedy, B. E. Hoppel, M. S. Cohen, and R. Turner *et al.*, "Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation," in *Proc. Nat. Academy of Sciences USA*, vol. 89, 1992, pp. 5675–5679.

[51] S. Ogawa, D. Tank, R. Menon, J. Ellermann, S. Kim, H. Merkle, and U. K, "Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging," in *Proc. Nat. Academy of Sciences USA*, vol. 89, 1992, pp. 51–55.

[52] M. J. McKeown, S. Makeig, G. G. Brown, T.-P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, "Analysis of fMRI data by blind separation into independent spatial components," *Hum. Br. Map.*, vol. 6, pp. 160–188, 1998.

[53] M. J. McKeown, T.-P. Jung, S. Makeig, G. Brown, S. S. Kindermann, T.-W. Lee, and T. J. Sejnowski, "Spatially independent activity patterns in functional MRI data during the Stroop color-naming task," *Proc. Nat. Acad. Sci. USA*, vol. 95, pp. 803–810, 1998.

[54] M. J. McKeown and T. J. Sejnowski, "Independent component analysis of fMRI data: Examining the assumptions," *Hum. Br. Map.*, vol. 6, pp. 368–372, 1998.

[55] M. J. McKeown, "Detection of consistently task-related activations in fMRI data with hybrid independent component analysis," *Neuroimage*, vol. V11, pp. 24–35, 2000.

[56] K. J. Friston, "Modes or models: A critique on independent component analysis for fMRI [Comment]," *Trends Cogn. Sci.*, vol. 2, pp. 373–375, 1998.

[57] M. J. McKeown, S. Makeig, G. G. Brown, T.-P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, "Response from McKeown, Makeig, Brown, Jung, Kindermann, Bell, and Sejnowski [Comment]," *Trends Cogn. Sci.*, vol. 2, p. 375, 1998.

[58] J. Kelso, A. Fuchs, R. Lancaster, T. Holroyd, D. Cheyne, and H. Weinberg, "Dynamic cortical activity in the human brain reveals motor equivalence," *Nature*, vol. 392, pp. 814–818, 1998.

[59] S. A. Berns GS and H. Mao, "Continuous functional magnetic resonance imaging reveals dynamic nonlinearities of 'dose-response' curves for finger opposition," *J. Neurosci.*, vol. 19, p. RC17, 1999.

[60] C. Moritz, V. Haughton, D. Cordes, M. Quigley, and M. Meyerand, "Whole-brain functional MR imaging activation from a finger-tapping task examined with independent component analysis," *Amer. J. Neuroradiol.*, vol. 21, pp. 1629–1635, 2000.

[61] K. Arfanakis, D. Cordes, V. Haughton, C. Moritz, M. Quigley, and Meyerand, "Combining independent component analysis and correlation analysis to probe interregional connectivity in fMRI task activation datasets," *Magn. Res. Imag.*, vol. 18, pp. 921–930, 2000.

[62] K. Petersen, L. Hansen, T. Kolenda, E. Rostrup, and S. Strother, "On the independent components of functional neuroimages," in *Int. Workshop on Independent Component Analysis (ICA 2000)*, 2000.

[63] S. Makeig, T.-P. Jung, D. Ghahremani, and T. J. Sejnowski, "Independent component analysis of simulated ERP data," in *Hum. High. Func. I: Adv. Meth.*, T. Nakada, Ed., 2000.

[64] I. Schiessbl, M. Stetter, J. E. Mayhew, M. N, J. S. Lund, and K. Obermayer, "Blind signal separation from optical imaging recordings with extended spatial decorrelation," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 573–577, 2000.

[65] M. Laubach, M. Shuler, and M. A. L. Nicolelis, "Independent component analyzes for quantifying neuronal ensemble interactions," *J. Neurosci. Meth.*, vol. 94, pp. 141–154, 1999.

[66] M. Laubach, J. Wessberg, and M. A. L. Nicolelis, "Cortical ensemble activity increasingly predicts behavioral outcomes during learning of a motor task," *Nature*, vol. 405, pp. 567–571, 2000.

[67] G. D. Brown, S. Yamada, and T. J. Sejnowski, "Independent component analysis (ICA) at the neural cocktail party," *Trends Neurosci.*, vol. 24, pp. 54–63, 2001.

[68] K. Prank, J. Borger, A. von zur Muhlen, G. Brabant, and C. Schofl, "Independent component analysis of intracellular calcium spike data," *Adv. Neural Inform. Process. Syst.*, vol. 11, pp. 931–937, 1999.

[69] L. Y. Kermadi I, A. Tempini, and E. M. Rouiller, "Effects of reversible inactivation of the supplementary motor area (SMA) on unimanual grasp and bimanual pull and grasp performance in monkeys," *Somatosens. Motor Res.*, vol. 14, pp. 268–280, 1997.

[70] T.-P. Jung, S. Makeig, J. Townsend, M. Westerfield, B. Hicks, E. Courchesne, and T. J. Sejnowski, "Single-trial ERP's during continuous fMRI scanning," *Soc. Neurosci. Abstr.*, vol. 25, p. 1389, 1999.

[71] K. J. Friston, P. Fletcher, O. Josephs, A. Holmes, M. D. Rugg, and R. Turner, "Event-related fMRI: Characterizing differential responses," *Neuroimage*, vol. 7, pp. 30–40, 1998.

Martin J. McKeown in an Assistant Professor of Medicine (Neurology) and Biomedical Engineering at Duke University, and one of the core faculty at Duke University's Brain Imaging and Analysis Center (BIAC). He is a U.S. board-certified Neurologist and an attending physician at the Durham Rehabilitation Institute. After completing undergraduate training in engineering physics at McMaster University in 1986, he attended medical school at the University of Toronto, graduating in 1990. He specialized in Neurology at the University of Western Ontario, becoming a fellow of the Royal College of Physicians of Canada in 1994. After completing a year of clinical electrophysiology training, he studied blind source separation techniques for fMRI analysis for three years at the Computational Neurobiology Lab, Salk Institute for Biological Studies. Since 1998, he has been on faculty at Duke University. His primary interests are in brain plasticity and rehabilitation. This work involves the application of ICA and other multivariate signal processing techniques to fMRI, EMG, and EEG to explore the brain and muscle activity related to motor recovery after stroke. Other work involves the application of ICA to the clinical EEG and fMRI to aid in the diagnosis, treatment and prognosis of epileptic populations.

Tzyy-Ping Jung (Member, IEEE) received the B.S. degree in electronics engineering from National Chiao Tung University, Taiwan, in 1984, and the M.S. and Ph.D. degrees in electrical engineering from The Ohio State University in 1989 and 1993, respectively.

He was a Research Associate at the National Research Council of the National Academy of Sciences and at the Computational Neurobiology Laboratory, The Salk Institute, San Diego, CA. In 1998, he became a Research Scientist in the Institute for Neural Computation of University of California, San Diego, where he heads the Neuroengineering Lab. His research interests are in the areas of biomedical signal processing, cognitive neuroscience, artificial neural networks, time-frequency analysis of human EEG and functional magnetic resonance imaging, and the development of neural human–system interfaces.

Scott Makeig received the Ph.D. degree in music psychobiology from the University of California, San Diego (UCSD), in 1985. His dissertation concerned the use of frequency domain measures of the auditory 40-Hz steady-state response for monitoring changes in central state. Thereafter, he expanded his study to considering the relation of changes in the whole EEG spectrum to changes in alertness, results that he and colleagues used to develop a method for objective real-time alertness monitoring. In 1996, he and colleagues published the first application of independent component analysis (ICA) to EEG data.

In 2000, he joined the staff of The Salk Institute, La Jolla, CA, and became a researcher in the Institute for Neural Computation of UCSD where he heads a computational neuroscience laboratory applying new signal processing methods to cognitive neuroscience.

Anthony J. (Tony) Bell received the M.A. degree in computer science and philosophy from the University of St. Andrews, Scotland, in 1987 and the Ph.D. degree in artificial intelligence from the Free University of Brussels, Belgium, in 1993.

Since 1990, he has been associated with the Computational Neurobiology Laboratory of the Salk Institute, San Diego, as a visitor or postdoctoral researcher, and has also worked at Interval Research, Palo Alto, CA.

Te-Won Lee (Member, IEEE) received the diploma degree in March 1995 and his Ph.D. degree in October 1997 (*summa cum laude*) in electrical engineering from the University of Technology, Berlin.

He is an Assistant Research Professor at the Institute for Neural Computation, University of California, San Diego, CA. He is also a research associate at the Salk Institute for Biological Studies, La Jolla, CA. He was a visiting graduate student at the Institute Nationale Polytechnique de Grenoble, the University of California, Berkeley, and Carnegie Mellon University. From 1995 to 1997, he was a Max-Planck Institute fellow and in 1999, he was a Visiting Professor at the Korea Advanced Institute of Science and Technology (KAIST). His research interests are in the areas of unsupervised learning algorithms, artificial neural networks and Bayesian probability theory with applications in signal and image processing.

Terrence J. Sejnowski (Fellow, IEEE) received the B.S. degree from the Case-Western Reserve University, the M.A. degree from Princeton University, and the Ph.D. degree from Princeton University in 1978, all in physics.

He is an Investigator with Howard Hughes Medical Institute and a Professor at the Salk Institute for Biological Studies, where he directs the Computational Neurobiology Laboratory. He is also Professor of Biology at the University of California, San Diego, where he is Director of the Institute for Neural Computation. In 1988, he founded *Neural Computation*, (MIT Press). The long-range goal of his research is to build linking principles from brain to behavior using computational models.