NeuroImage 49 (2010) 2401–2415

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Improving human brain mapping via joint inversion of brain electrodynamics and the BOLD signal

Kevin S. Brown [a,c], Stephanie Ortigue [b,c,d], Scott T. Grafton [b,c,d,*], Jean M. Carlson [a,c]

[a] Department of Physics, University of California, Santa Barbara, CA 93106, USA
[b] Department of Psychology, University of California, Santa Barbara, CA 93106, USA
[c] Institute for Collaborative Biotechnologies, University of California, Santa Barbara, CA 93106, USA
[d] UCSB Brain Imaging Center, University of California, Santa Barbara, CA 93106, USA

ARTICLE INFO

Article history:
Received 13 May 2009
Revised 23 September 2009
Accepted 6 October 2009
Available online 13 October 2009

ABSTRACT

We present several methods to improve the resolution of human brain mapping by combining information obtained from surface electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) of the same participants performing the same task in separate imaging sessions. As an initial step in our methods we used independent component analysis (ICA) to obtain task-related sources for both EEG and fMRI. We then used that information in an integrated cost function that attempts to match both data sources and trades goodness of fit in one regime for another. We compared the performance and drawbacks of each method in localizing sources for a dual visual evoked response experiment, and we contrasted the results of adding fMRI information to simple EEG-only inversion methods. We found that adding fMRI information in a variety of ways gives superior results to classical minimum norm source estimation. Our findings lead us to favor a method which attempts to match EEG scalp dynamics along with voxel power obtained from ICA-processed blood oxygenation level dependent (BOLD) data; this method of joint inversion enables us to treat the two data sources as symmetrically as possible.

Introduction

The rapid growth of large-scale, high-precision neuroimaging technology has allowed the study of cognitive processes formerly only accessible via behavioral measures. A key question in cognitive neuroscience is a fundamental one: when and where do task-related activations in the brain occur? Unfortunately, no single noninvasive imaging technology is sufficient to fully capture spatiotemporal brain dynamics on psychologically relevant temporal and spatial scales. High density surface EEG has millisecond temporal resolution, but it is based on 100–200 spatially correlated measurements on the scalp. Complementary to EEG measurements, mapping with fMRI has millimeter spatial resolution, but the low-pass filtering nature of the hemodynamic response function (HRF) makes it difficult or impossible to resolve dynamic events separated by less than several seconds.

Recently there has been increasing interest in combining spatially resolved BOLD signal measurements from fMRI with temporally resolved EEG data in order to enhance resolution in both space and time when trying to understand the neural basis of cognitive processes. Several methods have been proposed including Bayesian statistical methods (Phillips et al., 2005) and those based on linear time-invariant system theory (Liu and He, 2008). Work has also been done in regularizing magnetoencephalographic (MEG) data using fMRI measurements (Liu et al., 1998; Dale et al., 2000). The immediate challenge with such a joint approach is that the methods yield fundamentally different measurements in relation to underlying physiologic processes. A critical first step is to find a modeling framework that can represent both data types in a complementary way. In this paper, we do this by first modeling the two datasets with temporal and spatial ICA. From this a joint optimization that attempts to fit both EEG and fMRI data simultaneously can be developed. Independent component analysis (ICA) is a method for so-called blind source separation, and was originally developed in the context of speech signals (Jutten and Herault, 1991). ICA has been increasingly employed in medical imaging (McKeown et al., 1998; Makeig et al., 1996; Eichele et al., 2009) both for artifact removal in EEG recordings (Jung et al., 2000; LeVan et al., 2006; Ting et al., 2006) and to try to separate independent cognitive events. Investigators have used ICA on both fMRI and EEG data and attempted to combine the resulting information (Calhoun et al., 2006; Eichele et al., 2009, 2008; Moosmann et al., 2008). These joint decompositions either used EEG information from only a single electrode (Calhoun et al., 2006) or used the joint decomposition to attempt to avoid the EEG inverse problem

* Corresponding author. Department of Psychology, University of California, Santa Barbara, CA 93106, USA.
E-mail addresses: brown@physics.ucsb.edu (K.S. Brown), ortigue@psych.ucsb.edu (S. Ortigue), grafton@psych.ucsb.edu (S.T. Grafton), carlson@physics.ucsb.edu (J.M. Carlson).

1053-8119/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2009.10.011

(Moosmann et al., 2008). Additionally, they required truly simultaneous EEG and fMRI measurements, which will not be available to all investigators interested in combining the two imaging modalities.

We employ ICA as well, but emphasize that ICA alone does not solve the mapping resolution problem. ICA is a very useful data compression technique and in the context of neuroimaging experiments can be an invaluable "automatic" method for artifact removal (LeVan et al., 2006; Ting et al., 2006). We use ICA to clean and compress the data, but additional machinery is required to merge the two imaging modalities. We present a principled way of performing this merger, using joint cost functions that simultaneously match both EEG and fMRI data to obtain fine temporally and spatially resolved brain maps.

Our approach is to treat the determination of source dynamics in the brain as a large joint inverse problem; we build on and expand the work using synthetic data in Brookings et al. (2009). Sources of electrical activity in the cortex arising as a result of particular tasks give rise to both EEG signals, via volume conduction through brain, cerebrospinal fluid, skull, and scalp, and measured fMRI signals, via the BOLD response. Our methods generate a single set of dynamic brain sources that best reproduce the observed EEG and fMRI signals.

The EEG-only inverse problem is highly underdetermined (see below), and the way in which we should incorporate fMRI information is unclear. For this reason, we compare and contrast several different joint inverse problems that incorporate fMRI measurements in different ways. We consider both quadratic cost functions, which can be solved with straightforward linear algebraic techniques, and higher-order cost functions that require iterative methods. We show promising results for several of these methods in a dual visual evoked response experiment with human participants. We argue that experimental design and study goals are important considerations when deciding how to fuse EEG and fMRI to obtain high-resolution human brain maps.

Materials and methods

The data processing pipeline for joint inversion is illustrated schematically in Fig. 1. The steps in this pipeline are discussed in more detail in the following sections.

Experimental design

Participants

Twenty volunteers (five female) participated in the present study involving EEG, fMRI, or both. All were right-handed by the Edinburgh Handedness Inventory (Oldfield, 1971) and had normal or corrected-to-normal visual acuity. None had any prior or current neurological or psychiatric impairment, as ascertained by a detailed anamnesis. Mean age of participants was 24 years (range 18–44 years). Here we focus on data from four volunteers of the group who were studied with both EEG and fMRI. Prior to participation, volunteers provided written informed consent that had been approved by the Ethical Committee of the University of California, Santa Barbara.

Task and procedure

Participants observed a pair of static images obtained from videos of hands interacting with objects (see Fig. 2). In the first frame, an object was shown (a coffeemaker, a cup, cookies, etc.). This was followed by a second frame showing a hand grasping the same object or touching it. There were two types of familiar/meaningful grasping actions and the touching action. For the purpose of this experiment the different types of actions were collapsed into one sample. The presentation of the pairs of stimuli was interspersed with periods of blank screen including a fixation cross. Experimental blocks were intermixed for every participant.

Every trial consisted of the following sequence: a 150 ms fixation cross, a 500 ms single object, and a 2000 ms hand-on-object action. The same design was used in EEG and fMRI experiments. For the EEG experiment, intertrial interval (ITI) varied in 100 ms random increments from 1000 to 2000 ms. For the fMRI experiment, ITI varied from 2000 to 6000 ms in 2000 ms random increments. The stimulus presentation used a discrete trial procedure so that the temporal gap between the first and the second image of a video-clip was very short (one refreshed screen) in order to simulate a continuous image sequence that created a percept of an action.

Participants received explicit instructions to observe carefully all the stimuli, and to try to decode why a familiar/meaningful action was performed (modified version of the intention inference task; Ortigue et al., 2009). In order to avoid any saccadic movements, participants were asked to fixate on a central visual cross during the whole experiment. Before recording, every participant was familiarized with all actions for 3 min.

For the EEG study, visual stimuli were presented on a PC computer using Cogent 2000 (http://www.vislab.ucl.ac.uk/Cogent2000/index.html) running in Matlab 7.0.1 under Windows XP, which provides control of display durations and accurate recordings of reaction times. Participants were comfortably seated 150 cm away from a PC computer screen in which video clips were presented centrally. A total of five experimental blocks were presented throughout the whole experimental session. A total of 240 trials were administered which took up to a total of 40 min including breaks between each block.

During the fMRI recordings, visual stimuli were back-projected onto a screen at the head of the scanner bore by a standard LCD projector; participants viewed the screen using a mirror mounted on the head coil. Stimulus presentation was controlled with Psychtoolbox and Matlab 7.4.0 (R2007a) running on the Mac OS X platform. Functional images were collected in four discrete runs of 90 images each.

EEG data collection and preprocessing

Continuous surface electroencephalogram (EEG) was recorded from 128 AgCl carbon-fiber coated electrodes using an Electric Geodesic Sensor Net (GSN300; Electrical Geodesic, Inc., Oregon; http://www.egi.com/), where EEG electrodes are arrayed in a regular distribution across the head surface and the inter-sensor distance is approximately 3 cm. The EEG was digitized at 500 Hz (corresponding to a sample bin of 2 ms), band-width of 0.01–200 Hz, with the vertex electrode (Cz) serving as an on-line recording reference. Impedances were kept below 50 k. Data logging were via NetStation Recorder. Data were collected in multiple sessions with brief intervening rest periods for the participant. Before ICA was performed, each electrode was linearly detrended, the multiple sessions were concatenated, and the data were band pass filtered between 1 and 30 Hz.

Magnetic resonance imaging recordings and preprocessing

Both functional and electrical neuroimaging was conducted at the UCSB Brain Imaging Center. fMRI recordings were conducted using a 3T TIM Trio Siemens Magnetom with a 12 channel phased-array head coil. Foam padding was used for head stabilization. For each functional run, an echo planar gradient-echo imaging sequence sensitive to BOLD contrast was used to acquire 33 slices per repetition time (TR) (3 mm thickness, 0.5 mm gap), with a TR of 2000 ms, echo time (TE) of 30 ms, flip angle of 90°, field of view (FOV) of 192 mm, and 64×64 matrix. Before all the functional runs, a high-resolution T1-weighted mprage sequence image of the whole brain was acquired (TR = 2300 ms; TE = 2.98 ms; flip angle = 9°, 3-D acquisition, FOV = 256 mm; slice thickness = 1.1 mm, matrix = 256×256).

Initial data processing was performed with SPM 5.0 (http://www.fil.ion.ucl.ac.uk/spm/). fMRI image volumes were slice time corrected, motion corrected, unwarped, spatially normalized to the Montreal

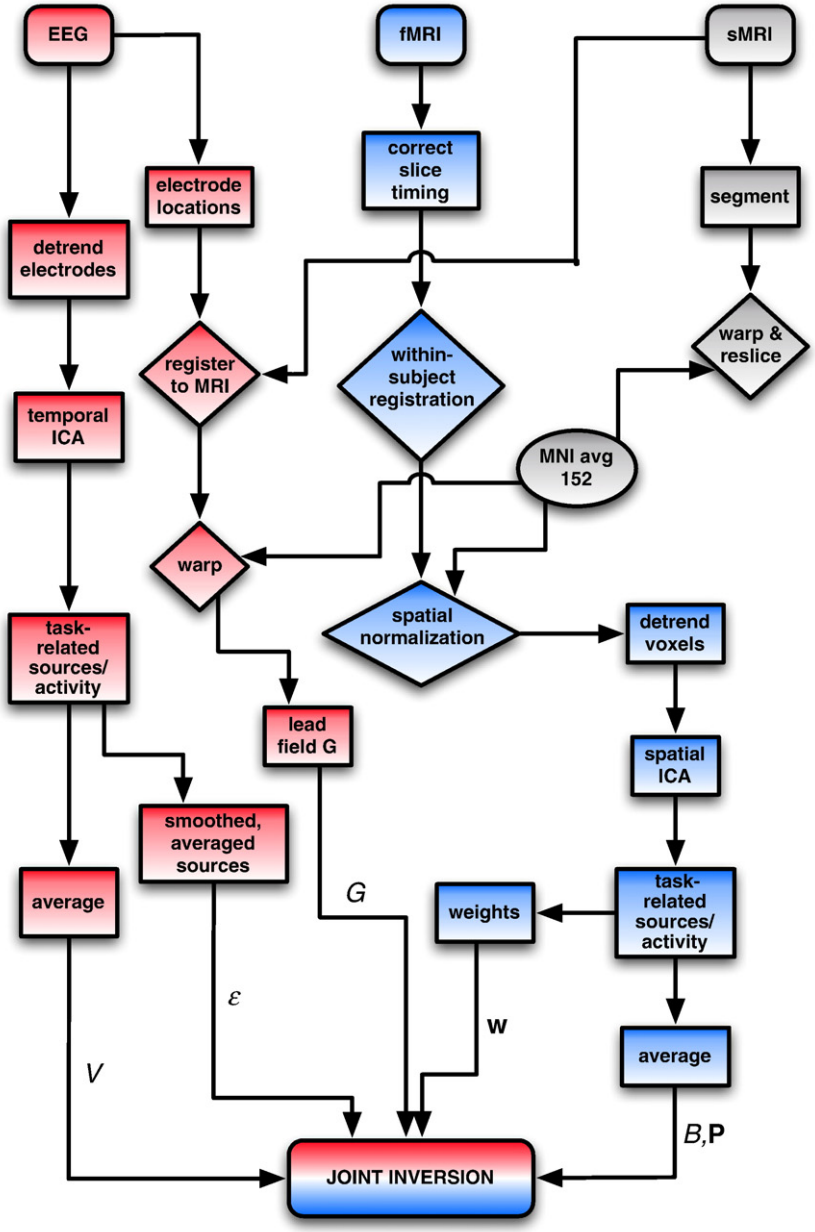


Fig. 1. Box diagram illustrating data pipeline steps leading to joint inversion. Consult the Materials and methods section for further details on these processing steps and the methods used in joint inversion.

152 Average T1 atlas, and resliced to 3 mm × 3 mm × 3 mm voxel sizes. Each voxel in every imaging session was linearly detrended and then the sessions were concatenated; these steps occurred after all SPM preprocessing. The high-resolution anatomical MRI was segmented into gray and white matter, warped to the 152 T1 atlas, and resliced to 3 mm × 3 mm × 3 mm voxel sizes in order to obtain a set of potential solution points. No spatial smoothing was performed on the functional images.

Independent component analysis

Independent component analysis was performed on both the EEG data and the fMRI data, in each case using projection pursuit as implemented in FastICA (Hyvarinen and Oja, 1997, 2000). Our decompositions are similar in spirit to those used by other investigators (Jung et al., 2001; McKeown et al., 1998) and we employ both temporal and spatial ICA (Makeig et al., 1996;Hyvarinen and Oja, 2000; McKeown et al., 1998; Stone, 2004). In the case of EEG data, the ICA decomposition is

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

where each row of E is the time series of one electrode; we decomposed the full experimental session before doing any epoching or averaging. In our notation A is the mixing matrix and s is the matrix of statistically independent sources; there need not be as many sources as electrodes, but we attempted to extract as many sources as possible, up to the number of electrodes. This is temporal ICA; the statistically independent sources that are extracted are time series. The rows of s are the independent sources that mix to give the EEG data, and the columns of A give the so-called scalp maps, which represent the degree to which a source is present in each electrode.

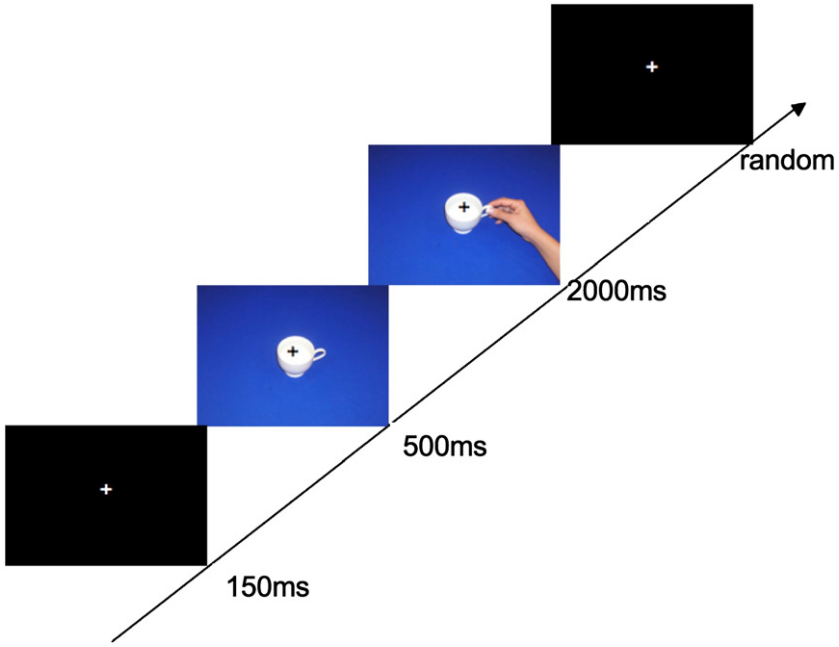


Fig. 2. Stimulus presentation for the dual visual evoked response experiment discussed in this manuscript.

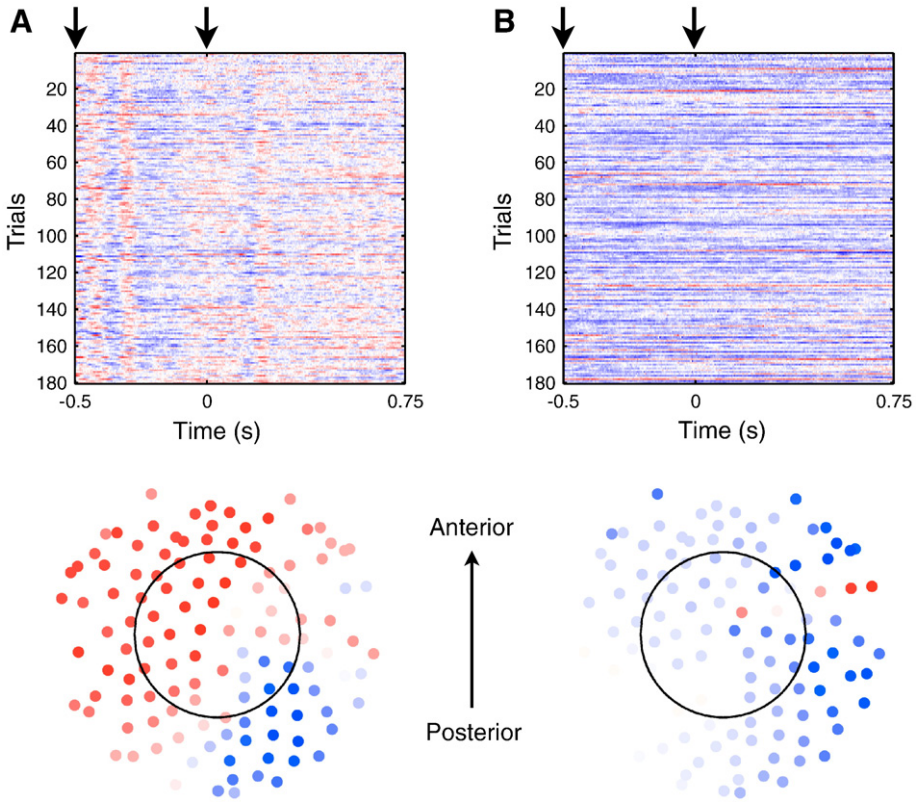


Fig. 3. Example of task- and nontask-related sources obtained from full-session ICA of human evoked visual response data (see Materials and methods). For both sources, we show a raster plot of the source, aligned with respect to stimulus presentation, and a Lambert equal-area polar projection of a spherical fit to electrode locations displaying the scalp map, obtained from the ICA mixing matrix (see Eq. 1). A solid black line is drawn in the scalp map to represent the equator of the fitting sphere. Stimuli were presented at −0.5 s and 0 s and are marked by the heavy black arrows in the raster plot. Anterior and posterior electrodes are indicated. In all panels red corresponds to positive and blue to negative. The left panel (A) shows a task-related component. There is strong trial-to-trial synchronization in the raster plot following stimulus presentation, and the scalp map is quite dipolar. The right panel (B) shows an artifactual or noise component; there is little consistency across trials, and the source has nothing to do with the stimulus. The scalp map is also quite unphysical; a few electrodes have large positive loading, but even their nearest neighbors show no activity or activity of the opposite sign rather than a smooth transition between positive and negative loading. We included the source in (A) and excluded the source in (B) in our analysis.

We use spatial ICA on the fMRI data; we write the decomposition as

$$F^T = DS \tag{2}$$

F^T is a matrix with the same number of rows as there are fMRI image volumes, and it has as many columns as there are voxels. Here, the sources S are independent spatial maps, and the corresponding column of D gives temporal variation of a map. In this case the sources S are statistically independent maps, not time series as in the EEG case (Stone, 2004).

Obtaining task-related sources

Many of the sources returned by both temporal and spatial ICA have nothing to do with the task; hence we must sift through the components to discover those that have something to do with the task. In the EEG case, we selected components by hand, as illustrated in Fig. 3. For each source (row of s in Eq. 1) returned by temporal ICA, we constructed an epoched raster plot, seen in the top row of Fig. 3. We also considered the corresponding scalp map (column of A in Eq. 1), shown in the bottom row of Fig. 3. Task-relevant sources showed strong vertical banding in the raster plot and a generally smooth scalp map. In Fig. 3, the source shown in panel A has these properties and that in panel B does not. All task-relevant sources—those that look like the source in panel A of Fig. 3—were chosen when rebuilding the EEG data (see below). Sources like those in panel B of Fig. 3 were not used as temporal basis functions in joint inversion and were not included in stimulus-locked time-averaged EEG data. The process of source selection could be made more quantitative and automatic, and we are currently working on metrics for picking EEG sources that would make the process more similar to our process for picking fMRI sources.

The selection of fMRI components was more automatic. A stimulus function $f(t)$, having the value 1 if an image was presented that TR and zero otherwise, was constructed and its correlation with $D(t)$ (see Eq. 2) was calculated for positive and negative lags between −20 and 20 s

$$C(\tau) = <D(t)f(t+\tau)>_t \tag{3}$$

The use of a lagged correlation in this case reflects the expected hemodynamic lag. The stimulus function was then randomly permuted 1000 times and lagged correlations recomputed; this gave a mean noise correlation and its standard deviation (denoted by the dotted blue lines in Fig. 4). Any component that showed above-noise correlations at negative lags, corresponding to the response following the stimulus, was then selected as task-related. In all cases for both fMRI and EEG data, the number of task-related sources was much smaller than the total number extracted (a few as opposed to hundreds), so ICA served as a useful data compression technique.

Epoching and basis function calculation

We used information obtained from ICA of EEG and fMRI data as follows, discussing the process for EEG data first. Supposing we found

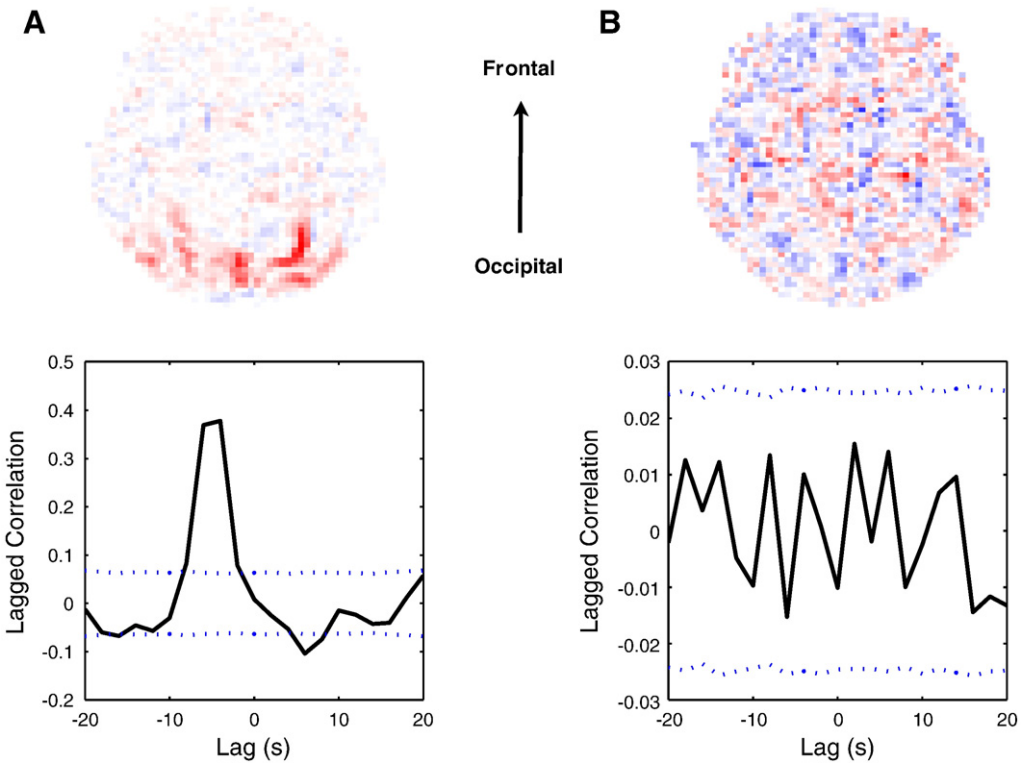


Fig. 4. Example of task and nontask-related sources obtained from full-session ICA of fMRI of human evoked visual response data (see Materials and methods). For each source we show a representative slice from the full spatial source, which extends to all gray and white matter voxels. We also show the lagged correlation of the temporal activity of that source —obtained from the mixing matrix (see Eq. 2)—with a binary function encoding stimulus presentation. In the correlation plot, positive lags correspond to fMRI activity leading image presentation and negative lags correspond to activity following presentation. The dotted blue lines give a significance level for the stimulus correlation, computed as described in Materials and methods. The left panel (A) shows a task-related component. Notice the head map has contiguous areas of strong activity, particularly near the rear of the head, as we expect for visual stimuli. The lagged correlation plot shows a strong correlation at 6–8 s after stimulus presentation. The right panel (B) displays a noise component; the head map looks essentially random, and there is no significant correlation with the stimulus at any lag. We included the source in (A) and excluded the source in (B) in our analysis.

m task-related sources, labeled {1,2,…,m}, we reconstructed the EEG voltage data as follows:

$$\tilde{E} = \begin{bmatrix} A_1^T A_2^T \ \dots \ A_m^T \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_m \end{bmatrix} \tag{4}$$

The s variables are row vectors of the task-related sources, and the A^T are column vectors of the corresponding scalp maps (stored in the matrix A, see Eq. 1). This operation gives a data matrix of the same size as E, but with artifacts and task-unrelated activity removed. To obtain the voltage matrix V which we used in all subsequent inverse methods, we then time-averaged $\tilde{E}$ locked to the stimulus of interest. This produced a matrix V with as many rows as there are electrodes, and as many columns as there are time samples in the size of the epoch we have chosen for averaging. For example, there would be 1000 time samples in 2 s of 500 Hz EEG data.

To obtain the matrix of basis functions ε used below (see the Inversion algorithms section), we time averaged the components $\mathbf{s}_i$, locked to stimulus presentation, and then low-pass filtered them with a filter width of 32 ms. This process ensures that we will avoid overfitting the noise in the EEG data; a linear combination of the smoothed sources cannot perfectly fit the data in V, whereas a direct ICA decomposition of V could reconstruct V with zero error. Our low-pass filter is one commonly used in geophysical research (Flagg et al., 1976; Limeburner et al., 1983); it attempts to balance frequency ringing from filters of finite (temporal) support with the reduction in peak amplitudes from infinite (temporal) support filters. For a width of $\tau_c = 1/f_c$, the filter equation is

$$f(t) = \frac{2\sin(2\pi f_c t) - \sin(\pi f_c t) - \sin(3\pi f_c t)}{\pi^2 f_c^2 t^3} \tag{5}$$

Simpler filters like Gaussian (infinite support) and Epanechikov (quadratic, finite support), applied with the same cutoff, give virtually indistinguishable final results.

Rebuilding the fMRI data proceeded in a manner very similar to that for EEG data; the task-related bold data F was built from task-related fMRI components as in Eq. (4), using the appropriate rows and columns of the D and S matrices in Eq. 2. Explicitly, if there were n task-related fMRI sources, labeled {1,2,…,n} we computed

$$\tilde{F}^T = \begin{bmatrix} D_1^T D_2^T \ D_n^T \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{bmatrix} \tag{6}$$

This data was then time averaged, again locked to stimulus, to obtain a matrix of BOLD data B. The matrix B has as many rows as there are fMRI voxels and as many columns as there are time samples in the epoch chosen for averaging (360 columns for this experiment). We also use the fMRI components to compute a set of spatial weights w, sometimes arranged for convenience in a diagonal matrix W (with w on the diagonal). To construct w we found all the spatial maps carrying task-related information and then formed $w = \sum_{j=1}^{n} |S_j|$. We subsequently normalized w so that its largest element equals unity. As we will soon show ("Inversion algorithms," below), we sometimes use w in our joint inversion methods and sometimes B. When employing w, we are ignoring any voxelwise temporal information from fMRI and using only the spatial map of task-related activity. When we incorporate fMRI data as, B we are including the temporal dynamics of each voxel, and are hence trying to fit not only a spatial activity map but the fMRI dynamics at the voxel level.

There is a symmetry in the way the fMRI and EEG results are treated; in each case the mixing elements were used in selecting task-related components, but then that information was discarded, and only the sources were kept. In the case of fMRI, all task-related sources were collapsed into one vector w, rather than maintaining them separately as in the EEG case. If the experiment were truly recorded simultaneously (EEG measurements inside an MRI scanner) we might be able to associate EEG sources with fMRI sources, but we have no confidence in our ability to do so when the imaging is performed in separate sessions, nor do we necessarily expect such a correspondence in this situation.

Lead field calculation

Electrode locations were measured using an infrared tracking system (Northern Digital), and put in registry with the participant's structural MRI by using anatomical fiducials also measured with the TMS tracker (outer canthus of the eyes and the tip of the nose). These electrode positions were then warped into template coordinates using SPM 5, and the subsequent warped positions used for all subjects. Three-shell spherical lead fields were calculated using Berg's method (Berg and Scherg, 1994), employing the BrainStorm toolbox for MATLAB (http://neuroimage.usc.edu/brainstorm/). Each shell was assumed to have a different isotropic, uniform conductance. Gain matrices obtained from lead field calculation were depth weighted (using electrode power) to discount superficial sources (Kohler et al., 1996).

A subset of gray matter voxels obtained from anatomical image segmentation was chosen to be solution points. Only voxels with gray matter intensity above a threshold of 75% of maximum were chosen, in order to restrict the size of the solution grid. The size of the solution grid was further decreased by selecting a desired number at random. Dipole orientations were chosen to be radially outward from the center of the head; hence all solution methods yield a scalar (rather than a three-component vector) at each solution point, and the lead field gain matrix reflected this choice of dipole orientation.

Inversion algorithms

Inferring which sources in the brain produced measured EEG signals is a problem from classical electrodynamics; we wish to obtain the dynamic charge distribution in the interior of a set of roughly spherical conducting shells using only measurements on the surface of the outermost shell. Given that classical electrodynamics is a well-understood theory, this seems straightforward but is hardly so. One can represent the forward problem as

$$V = GX \tag{7}$$

where V is the matrix of electrode voltages (all electrodes, all times), G is the linear Green's function which propagates the interior sources to the scalp, taking account of geometry and material properties of cortex, cerebrospinal fluid, skull, and scalp, and X is the desired matrix of sources (Hallez et al., 2007).

Formulated as a least-squares problem, the desired matrix of sources minimizes the following cost function:

$$f_{EEG}(X) = \frac{1}{2}\mathrm{Tr}\left\{(V - GX)^T (V - GX)\right\} \tag{8}$$

G has a huge null space; we would like source information over the entire cortex—tens of thousands of locations, depending on desired resolution—at all times, but given typical EEG electrode montages we have at most around 200 correlated spatial measurements to constrain the solution. While anatomical information (Phillips et al., 2002) or careful consideration of the types of bulk currents that produce EEG signals (de Peralta Menendez et al., 2000) can further constrain the problem, all EEG-only mapping methods must eventually impose regularization constraints to give a unique X. These constraints are

usually chosen to impose desirable properties on the solution, like that it be of minimum norm (Hämäläinen and Ilmoniemi, 1994) or maximum smoothness (Pascual-Marqui et al., 1994), or have a particular covariance structure (Sekihara et al., 2001). For a recent review of EEG-only source localization see Grech et al. (2008).

We incorporate fMRI information because it has complementary strengths and weaknesses to EEG; it has high spatial resolution but very poor temporal resolution, on the order of several seconds. The single term in Eq. 8 fits only the observed EEG electrode voltages V, and we will always represent the EEG portion of the cost function this way. In order to incorporate fMRI information we will modify the cost function in Eq. 8 by adding terms constructed to fit fMRI data. There are many ways to do this, and comparing and contrasting some of them are the focus of the rest of this manuscript.

We summarize the joint cost function approach with the following "master" cost function, which we wish to minimize:

$$C(X) = \frac{1}{2}\mathrm{Tr}\left\{(V-GX)^T\Sigma_V^{-1}(V-GX)\right\} + \frac{\mu_1}{2}f_{fMRI}(X) + \frac{\mu_2}{2}f_R(X) \quad (9)$$

The first term represents the fit to the EEG data V, and it is identical to Eq. 8 except for the addition of Σ_V, discussed below. The second term is the fit to fMRI data, and the third is a regularizing term. These final two terms are written less explicitly than the EEG term because they will have different representations in different joint inversion methods. However, whatever form $f_R(X)$ and $f_{\mathrm{fMRI}}(X)$ take, both must be a function of the brain sources X (these are not the ICA sources) in order to relate the two imaging modalities and best constrain the parameters X using all available experimental measurements.

Σ_V is the electrode–electrode covariance matrix. Scaling by this matrix serves two purposes; the off-diagonal elements ensure that we do not overweight correlated data points, and the on-diagonal terms ensure the relative contributions to the cost function of terms with very different units are similar. We use the full-session EEG data, before epoching and averaging but after filtering/rebuilding by ICA (Eq. 4), to compute the electrode–electrode covariances in Σ_V. The parameters μ_1 and μ_2 control the relative weights of the various terms; we include them for completeness, but we generally only consider equal weighting ($\mu_1=\mu_2=1$) in this manuscript. Detailed descriptions of each method follow, and a table summarizes them (see Table 1).

Minimum Norm (MN)

Minimum norm is a well-known method for regularizing under-determined least-squares problems (Wunsch, 1996). It is a widely used method to solve the EEG-only source localization problem (Hämäläinen and Ilmoniemi, 1994). Minimum norm simply seeks a solution in which the source matrix X is as small as possible, in the sense of having minimum L_2 norm. The second term in Eq. 9 is not present, and the regularization term becomes

$$f_R(X) = \mathrm{Tr}\left\{X^TX\right\} \quad (10)$$

which is the correct expression for the sum of the squares of the elements of the matrix X. Despite, or perhaps because of, its simplicity, some investigators continue to argue for the use of minimum norm estimation (Hauk, 2004).

ICA Only (ICAO)

This is a modification of the minimum norm procedure. Instead of solving for the full matrix X, which has row dimension equal to the number of fMRI voxels and time dimension equal to the number of sampled EEG voltages, we posit that all source activities can be written as a linear combination of our previously described ICA-derived EEG basis functions, where

$$X = \alpha\varepsilon \quad (11)$$

with ε the averaged, smoothed, task-related EEG sources and α a matrix whose column dimension is much smaller than the number of times at which EEG data is collected. We refer to α as the loadings, as they are static spatial weights. The temporal dynamics of the solution are contained in ε. In any method employing this decomposition, all the functions in Eq. 9 become functions of α instead of X.

The first term in Eq. 9 may be written as

$$f_{EEG}(\alpha) = \frac{1}{2}\mathrm{Tr}\left\{(V-G\alpha\varepsilon)^T\Sigma_V^{-1}(V-G\alpha\varepsilon)\right\} \quad (12)$$

This equation is implemented here and in all subsequent methods which use ICA decomposition of the source activity. The regularization term in this case may be written as $f_R(\alpha)=\mathrm{Tr}\{\alpha^T\alpha\}$; it is minimum norm regularization but on the reduced rank representation α rather than X. We potentially gain something with this technique even though it is still EEG only; the use of ICA to generate a set of independent basis functions allows us to spatially separate the dynamics of independent cognitive processes, insofar as ICA is able to detect them in the EEG data. We also reduce the number of parameters we need to determine by a factor of 100 or more—the column dimension of α is of order 10, rather than 10^3 as with X.

W-scaled Gain Matrix (WGM)

In this method, we column-scale the lead field by the diagonal matrix W obtained from ICA of fMRI. This has the effect of suppressing solution intensity from solution points where no significant task-related fMRI activity was found. Mathematically, there is still no direct fMRI term, but a modified gain matrix $G=GW$ is used for inversion. Replacing G by G in Eq. 12 gives the EEG fitting term for this method. We again include a minimum norm term on α, so $f_R(\alpha)=\mathrm{Tr}\{\alpha^T\alpha\}$.

Table 1
A summary and quick reference for the inverse methods and their abbreviations used in the manuscript and figures; for each method, information about each of the three terms in Eq. 9 is given, along with the parameters determined in each optimization.

Method name (abbreviation)	Parameters	f_{EEG}	f_{fMRI}	f_R
Minimum Norm (MN)	X (~10^7)	Full solution matrix	None	L_2 norm on X
ICA Only (ICAO)	α (~10^5)	Loadings	None	L_2 norm on α
W-scaled Gain Matrix (WGM)	α	Loadings, G scaled by fMRI	None	L_2 norm on α
W_2-regularized ICA (W2ICA)	α	Loadings	None	L_2 norm of fMRI-scaled solution
W_1-regularized ICA (W1ICA)	α	loadings	None	L_1 norm of fMRI-scaled solution
Power Constrained (PowR)	α, scalar λ	Loadings	Solution power fit to fMRI power at each voxel	None
Model Reduced Joint Inverse (MRJI)	α, scalars λ, κ	Loadings	fMRI dynamics fit to balloon model basis β	None

G is the lead field matrix (see Eq. 7). X and α (see Eq. 11) both have a number of rows equal to the number of MRI voxels used in performing the joint inverse. X has a number of columns equal to the number of EEG sampling times, and α has a number of columns equal to the number of task-related sources in the EEG data. In calculating order-of-magnitude sizes for X and α we have assumed a solution grid of size 10000 voxels, 1000 EEG sampling times, and 10 task-related EEG sources.

W_2-regularized ICA (W2ICA)

This method still has no direct fMRI term; fMRI information is included in the regularization term. The form of the third term in Eq. 9 is

$$f_R(\alpha) = \mathrm{Tr}\left\{\alpha^T W^{-1}\alpha\right\} \tag{13}$$

To understand the purpose of this term, note that it is similar to the minimum norm constraint in Eq. 10 except that the loadings α (rather than the full source activity) are being constrained and W^{-1} is present. Multiplication by W^{-1} has the effect of increasing the penalty term for solution points which show little task-related activity (small W_{ii}, hence large $1/W_{ii}$). The less fMRI activity found at a solution point, the bigger a penalty this term introduces, and hence the more its activity will be suppressed in the inverse solution.

We should note here that ICAO, W2ICA, and WGM are all members of a family of cost functions; they all can be thought of as having an EEG fitting term like that in Eq. 12, and a regularizing term

$$f_R(\alpha) = \mathrm{Tr}\left\{\alpha^T W^{-p}\alpha\right\}. \tag{14}$$

ICAO has $p=0$ (thereby completely ignoring W), since its regularizing term is $\mathrm{Tr}\{\alpha^T\alpha\}$. For WGM, the transformation $\alpha\rightarrow W\alpha$ yields an EEG fitting term identical to Eq. 12 but with a regularizing term of $\mathrm{Tr}\{\alpha^T W^{-2}\alpha\}$. As p gets larger, the solution is penalized more heavily for placing solution power at voxels where fMRI activity is small; as $p\rightarrow\infty$ only the single most active fMRI voxel would be allowed to have solution activity. Hence, WGM forces more stringent correspondence with fMRI activity than W2ICA. (We thank an anonymous reviewer for pointing out this correspondence.)

W_1-regularized ICA (W1ICA)

This method is very similar to W2ICA except for the form of the regularization term

$$f_R(\alpha) = |W^{-1}\alpha| \tag{15}$$

which uses the L_1 norm rather than L_2. The L_1 norm has been used with great success in many regularized optimization problems, and in Bayesian statistical problems would correspond to a Laplacian rather than the more common Gaussian prior.

Operationally, the L_1 norm is often used for finding sparse solutions to inverse problems because the force on the parameters (derivative of the cost function) coming from the regularization term is constant rather than linear with X as in the L_2 case. No matter how small a parameter becomes, the L_1 term continues to push it towards zero until it is identically zero. L_2 terms become less and less effective as the parameters shrink, so it is difficult to obtain truly sparse solutions with them. Solving this optimization problem requires an iterative (though deterministic) algorithm for a matrix of parameters based on that previously described for a vector of parameters (Alliney and Ruzinsky, 1994). At a significant memory cost this minimization problem can also be formulated as a quadratic program. For recent applications of L_1 regularization in solving the EEG/MEG-only source localization problem, see Huang et al. (2006) and Ou et al. (2009).

Power Constrained (PowR)

This is the first of two methods which explicitly fit fMRI data, utilizing the second term in Eq. 9. This fMRI fitting term takes the form

$$f_{fMRI}(\alpha) = \frac{1}{\sigma_P^2}(\lambda P - P(\alpha))^T(\lambda P - P(\alpha)) \tag{16}$$

The vector $\mathbf{P}$ is the observed fMRI signal power at each solution point; its i^{th} element is computed from the reconstructed BOLD data B as

$$P_i = \sum_t B_{it}^2 \tag{17}$$

$\mathbf{P}(\alpha)$ is the calculated solution power, which for parameters α is

$$P_i(\alpha) = \left[\alpha C\alpha^T\right]_{ii} \tag{18}$$

where $C=\varepsilon\varepsilon^T$. The matrix C would be diagonal if ε came directly from ICA of the averaged EEG data (due to independence of the components), but our full-session EEG ICA and smoothing process introduces off-diagonal elements. The number σ_P^2 is the variance of the distribution of voxel power, which gives an appropriate scaling for the fMRI term relative to the EEG fitting term. The scalar λ absorbs the unknown conversion between BOLD units and solution intensity (essentially local field potential); it can be computed by setting $\partial C/\partial\lambda = 0$ and does not require a separate minimization step.

We seek to match scalp dynamics via EEG, and dynamics in the bulk brain by comparing solution power to observed voxel power from fMRI measurements. This algorithm must be solved via an iterative method; we use Møller's scaled conjugate gradient algorithm (Møller, 1993). While the cost is no longer simply quadratic, it is not highly nonlinear—merely quartic. Also, an analytical gradient is available making derivative-based methods accurate and easy to implement.

Model Reduced Joint Inverse (MRJI)

This method is described in more detail elsewhere (Brookings et al., 2009), but we summarize it here. The cost function contains both EEG and fMRI terms and is

$$C(\alpha,\lambda,\kappa) = \frac{1}{2}\mathrm{Tr}\left\{(V-G\alpha\varepsilon)^T(V-G\alpha\varepsilon)\right\} + \frac{g}{2}\mathrm{Tr}\left\{(\lambda B+\kappa-\alpha\beta)^T(\lambda B+\kappa-\alpha\beta)\right\} \tag{19}$$

This is the entire cost function; it replaces the first two terms of Eq. 9 and contains no third $f_R(\alpha)$ term. The goal is to treat fMRI and EEG as symmetrically as possible; both types of data are fit to dynamical models. $\beta=\beta(\varepsilon)$ is a matrix of fMRI sources, analogous to the EEG sources. These are obtained by feeding each EEG source (row of ε) into a balloon-type (Buxton et al., 1998) model for BOLD dynamics (Robinson et al., 2006). The β act as basis functions for attempting to fit the BOLD activity B in the way ε are used to fit the EEG dynamics. So in the same way that we assume the local source activity can, via propagation of the lead field, describe measured EEG voltage, we assume that fMRI activity can be written as a linear combination of fMRI basis functions, corresponding to computed BOLD responses to the sources ε.

The details of the solution in this case and the performance of the method on synthetic data are contained in Brookings et al. (2009). This cost function is entirely quadratic in the parameters (λ, α and κ) allowing us to find a (global) minimum using only linear algebra. g is a gain term that attempts to make the EEG and fMRI terms in the cost function of the same magnitude; we use the ratio of the average electrode variance to the average voxel variance. Note there is no covariance weighting in the function; see the Discussion section for further comments on this issue.

Considered in total, the inversion methods detailed above form a sort of hierarchy. ICAO attempts to improve on MN by employing a reduced rank representation for the parameters (α rather than X). WGM and W2ICA both incorporate fMRI information as an L_2 regularizing term, but WGM imposes a stricter correspondence with

fMRI activity. W1ICA uses an L_1 regularizing term, in an effort to force voxels with very small solution activity to zero activity. PowR and MRJI fit successively more of the fMRI dynamics, with MRJI attempting to fit the dynamical data for voxels (fMRI) as well as electrodes (EEG).

Results

In this section we compare results for the inversion methods described above. We have chosen the visual evoked response to test our methods because it generates a strong cortical signal and much is already known about its dynamics and localization. To wit, we expect up to three visual evoked potentials—an early peak and a late single or double peak—appearing in the occipital cortex due to a visual stimulus. The reader should keep this in mind when considering the results presented below.

Figs. 3 and 4 show implementation of ICA on full-session human EEG (Fig. 3) and fMRI (Fig. 4) data and give examples of which components were included in mapping and which were not. These figures are discussed in detail in the "Materials and methods" section and we refer the reader to that section for details about our filtering by ICA.

Figs. 5–7 show the results of the inversion methods described in Materials and methods for the evoked visual response data. Fig. 5 shows a set of summary plots for each algorithm, for one of the four participants shown in Fig. 7. Identical information for the other three participants is shown in three supplemental figures. The first panel indicates the ability of each inversion method to reproduce observed EEG voltages; hence rows of the matrix V are plotted simultaneously with rows of the matrix GX (for the MN method) or the matrix $G\alpha\varepsilon$ (for all other methods, see Eqs. 8 and 11). The second panel shows the agreement between observed fMRI activity, in the form of the weight vector $\mathbf{w}$ obtained from ICA of the BOLD data, and the square root of $\mathbf{P}(\alpha)$, the power computed at each solution voxel (see Eq. 18).

For the MN method, X is not decomposed as in Eq. 11, so the solution power in this one case is computed exactly as in Eq. 17, except with the matrix X in the place of B. The final panel shows a representative slice with solution power superimposed on subject anatomy. The size of each circle corresponds to solution power at that voxel, and voxels with very little solution power have been omitted for the sake of clarity. These anatomical plots indicate if significant solution power has been placed in physiologically reasonable locations. In our subsequent discussion we focus on Fig. 5 as exemplary, but our statements about method performance are consistent across the four participants (see supplemental figures).

There are several notable features of Fig. 5. First, by scanning the first column one can see that all the inversion methods used fit the EEG data quite well. This is true for the entire head and not just representative electrodes. One can obtain a global goodness of EEG fit by looking at the EEG portion of the cost function at the best parameters. The MRJI fits almost perfectly due to lack of covariance weighting in the EEG fitting term; we discuss this below. W1ICA fits the EEG data slightly worse than the other methods, particularly at electrode T5 (in three of the four participants). More striking is that similar EEG quality of fit is obtained for drastically different spatial distributions of solution power, as one can see by examining columns two and three. In the extreme, for all four participants W1ICA was able to obtain a respectable fit to the EEG data by placing solution intensity at less than 10 of 4000 voxels! This extreme sparseness is a phenomenon that has been observed in other L_1 mapping approaches, and the spatial discontinuities we see (highly active voxels in direct proximity to those of no activity) are similar to temporal "spikiness" obtained with L_1 methods applied at each timestep (Huang et al., 2006; Ou et al., 2009). These drawbacks have led other researchers to modify L_1 approaches using information from MEG physics (Huang et al., 2006) or to use L_1 regularization in the spatial domain and L_2 in the time domain (Ou et al., 2009). It would be interesting to see if similar approaches could be integrated into our framework in order to improve the performance of W1ICA.

One should also note that the WGM, W2ICA, and PowR methods all placed significant power in physiologically relevant regions. These results illustrate the underdetermined nature of the EEG-only problem and suggest there may be no single correct way to increase mapping resolution. We argue below that one needs to think about the interplay between joint inversion, experimental design parameters, and robustness issues when combining multiple imaging modalities for improved resolution in whole-brain mapping.

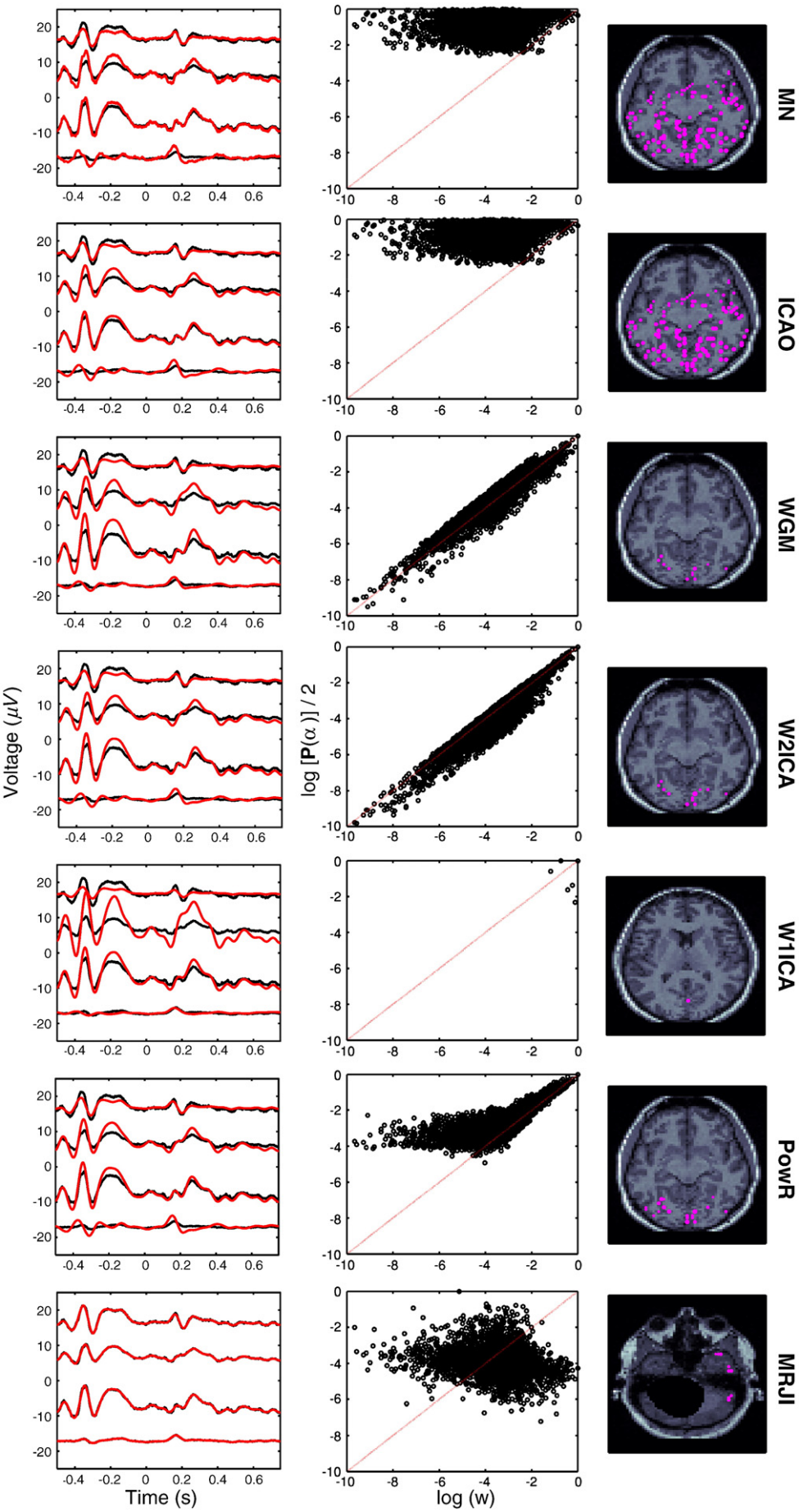
Fig. 6 shows the solution, rather than its diagnostics, for the subject shown in Fig. 5 for four of the methods we have discussed. Minimum norm has been left out because we here show α (the spatial loadings) and ε (the temporal signals) separately and basic minimum norm admits no such decomposition as part of its solution process. The ICAO method is essentially pre-decomposed minimum norm and is shown. We also do not display the MRJI or W1ICA, as their results did not look promising for any of the four subjects. For each method in Fig. 6 we show the individual rows of ε along with corresponding weights α, plotted as a scatterplot on subject anatomy. Negative loadings are in cyan and positive loadings in magenta, and a consistent scale for point size is used across all panels.

Notice that the loading maps are relatively smooth, despite having no explicit smoothness constraint in the cost functions. All the solution smoothness in WGM, W2ICA, and PowR comes directly from the smoothness of the activity map obtained from fMRI. This is a particularly nice feature, as the solution smoothness found is a side effect of fitting both modalities well, rather than a condition imposed by fiat. We also note that all three components used in the task-only EEG average (the top row of Fig. 6) have similar but not identical dynamical structure and localization. This is expected; the dominant response to the two successive stimuli will be the evoked visual response, and we do not expect nor desire that ICA will oversplit the paired visual stimulus into separate components for each image. However, for more complicated experimental paradigms in which we expect independent responses—for example visual evoked response followed by a motor response—ICA should be able to produce one component for each which, upon joint inversion, will show distinct, physiologically relevant localization.

Fig. 7 shows the temporal and spatial characteristics of the responses of the four study participants when computed using the PowR joint inversion method. First notice that the number of basis functions (rows of ε) comprising the solution varied from subject to subject. There is no guarantee that ICA will find the same number of task-related EEG components in different individuals, and we see that this is the case here. This statement is also true of spatial ICA of fMRI data; while all components for each participant have been collapsed into a single spatial map (not shown), the number of task-correlated components was variable among individuals. Fig. 7 shows intersubject differences in both activity timing and localization. However, common to all the subjects are single and double peaks following stimulus presentation, with significant occipital loading. These results show that joint inversion could be a powerful tool for probing intersubject variability, and in the "Discussion" section we elaborate on how one could begin to do so.

Discussion

We now compare and contrast the various inversion schemes presented. First, it is clear that ICA of EEG data alone does very little to condition the inverse problem; while it is valuable to pre-split temporally independent neural events before inversion mapping, the regularization problem remains, and minimum norm solutions produce substantial power at most solution voxels. Among the remaining methods, W1ICA is an interesting case, in that it was able



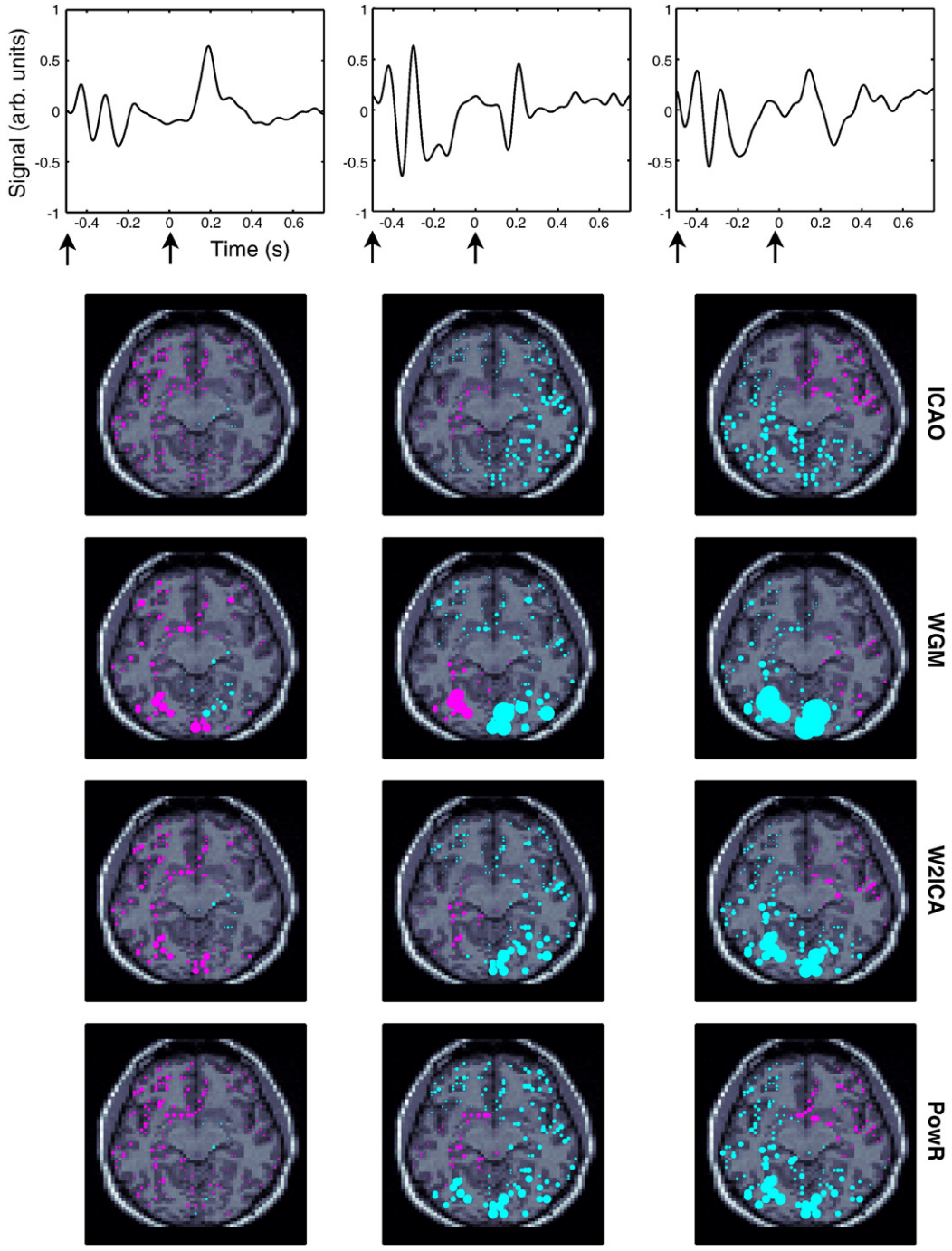
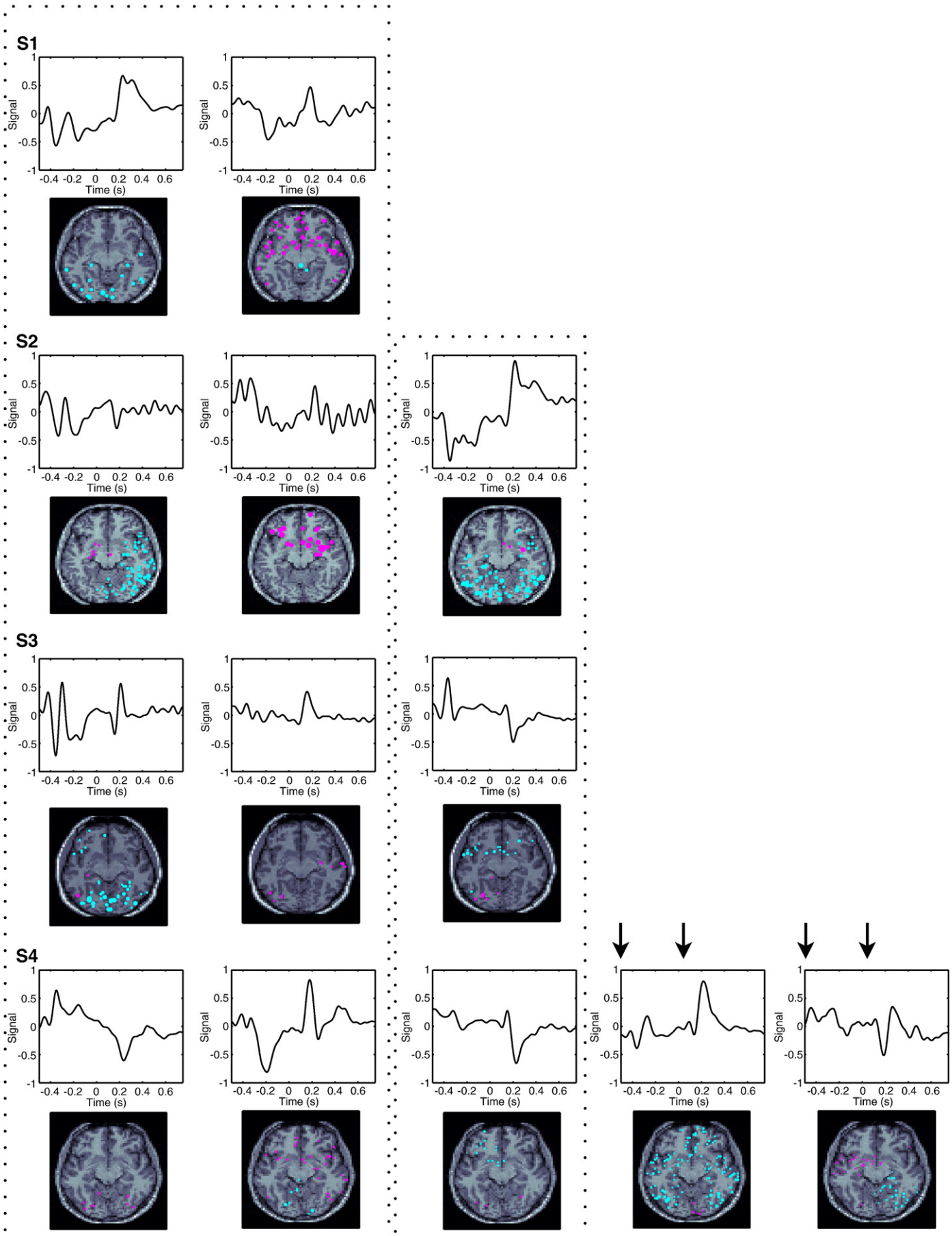


Fig. 6. Joint inverse solution for four of the methods described in the text: ICAO, WGM, W2ICA, and PowR, for subject S3 (shown also in Fig. 5). The top row of the plot shows the rows of ε, the matrix of temporal basis functions; heavy black arrows indicate stimulus presentation times. In the anatomical plot negative loadings (elements of α, the parameters obtained by inversion, see Eq. 11) are shown in cyan and positive loadings in magenta. The same slice is shown for each signal and all methods. To obtain the full solution for all solution voxels and all times, one constructs $\alpha\varepsilon$, which essentially multiplies the signal in the first row of the figure by the head map below it and then sums over the figure row.

Fig. 5. Solution diagnostics for the seven inversion methods described in the text, for one participant (S3 in Fig. 7) in the visual evoked response experiment. For each method, three panels are shown. The left column displays the quality of fit to the EEG data; the inverse solution is shown in red and the data in black, for four electrodes. They are, from top to bottom in the 10/20 naming system, P4, T5, O1, and F8. The voltage traces have been offset to allow for easier viewing. The second panel is a log-log plot of **w** (weights obtained from ICA of fMRI as described in the text) on the *x* axis and the square root of voxel power calculated from the inverse solution on the *y* axis. Hence we are plotting the power in the (fMRI) data (the *x* axis) versus the power in the model (the *y* axis). The red dotted line has unit slope; the amount of point scatter about this line shows the degree of disagreement with measured fMRI voxel power. The third column shows a representative anatomical slice with dots with size proportional to solution power at that voxel. In these plots, voxels with power less than 25% of the most powerful voxel have been omitted for clarity. Whenever possible we have chosen the same slice for all methods, but in the case of the MRJI and W1ICA that was not possible, as very little if any solution power was present in the slice shown for the other methods. Method abbreviations are indicated at right, and a quick reference is contained in Table 1. The same information for the other three subjects in Fig. 7 is shown in the supplemental figures.



to accurately reproduce the scalp voltage signal in all four subjects by putting power at only a few ($O(10)$) voxels; all other solution voxels had identically zero activity. One way to view this is as a side effect of basis construction; while we do smooth the averaged components from ICA to construct 6, the smoothing process is very mild, and therefore the individual basis functions still strongly resemble the signals from the electrodes. This would not be the case in a more complicated experiment with multiple different stimuli; each component might then be expected to encode only one of many independent event-related potentials.

Also too extreme in our estimation is WGM, which as we described in "Materials and methods" is very similar to W2ICA but with a stiffer penalty for placing solution activity at voxels with little fMRI activity. Given that both W2ICA and WGM place significant activity in physiologically relevant regions, we would opt for the softer penalty due to signal-to-noise issues in the fMRI data. We should point out, however, that p in Eq. 14 could also be included as a hyperparameter in the data pipeline and optimized rather than chosen a priori. Extensive simulations have previously shown the utility of L_2 regularization with fMRI (Liu et al., 1998), and this utility seems to carry over into our inversion framework.

MRJI was somewhat of a surprise. MRJI performed extremely well in tests on synthetic data (Brookings et al., 2009), but was far less successful with human subject data. This underperformance comes about for several reasons. For one, the individual rows of β are really all identical in this case; the low-pass filtering nature of the hemodynamic response makes events separated by less than several seconds unresolvable when used to force a Balloon/Windkessel model. In essence, rather than using a sum of distinct basis functions to fit the data in B, we are really only fitting a rescaled hemodynamic response function (HRF) to the time-locked fMRI average data. If the time-averaged voxel dynamics deviate from this simple shape, the fMRI term in the MRJI will not model them well.

This brings up the crucial subject of the interaction of inversion method with experimental design; this fMRI experiment was a rapid event-related paradigm, in which the responses to successive stimuli "pile up" in the measured fMRI data. Upon averaging, we do not recover simple rescalings of the HRF in the data. However, in a slower event-related design, where the hemodynamic signal is allowed to return to baseline after each stimulus presentation, this may not be true and the MRJI may approach its performance on simulated data (in which the assumed forward models generated the data, so a better fit was essentially ensured). It is therefore extremely important to consider experimental design when choosing how to improve mapping resolution using joint inversion; this is an issue which we are addressing with ongoing empirical studies.

Lack of similarity between the fMRI model employed and the observed data is not the only complication for the MRJI. Another comes when we try to properly weight the two terms in the cost function. Covariance weighting in least-squares cost functions performs two essential functions: it ensures we do not give ourselves more credit than we should when we fit correlated data, and it introduces a scaling term that can accommodate data in drastically different units, since the standard deviation should generally be of the same order of the mean and at least has the same units. While it is simple to include electrode–electrode covariance weighting in the MRJI, voxel–voxel covariance weighting poses a challenge. We cannot simply insert Σ_B^{-1}, the full voxel–voxel covariance matrix, into Eq. 19 in the obvious place; this is because of the parameter λ. Note that as written, λ rescales the data B to best match the model, for any choice of α. However, rescaling the data by λ rescales Σ_B^{-1} by λ^{-2}, and doing so destroys our ability to use the linear algebraic methods described previously (Brookings et al., 2009). For this reason, one would like to leave B alone and rather rescale $\alpha\beta$, which would adjust the units on the model to best fit the experimental units in the data. This again destroys our ability to use linear algebraic methods to solve the problem, as there are now terms quartic in the parameters present in the cost function ($\lambda^2\alpha^T\alpha$). We have used iterative methods to attempt to solve the problem in this case, but they prove particularly inefficient, partly owing to the fact that $a \equiv 0$ is not a valid initial starting solution—λ becomes undefined in this case. While the use of g in Eq. 19 tries to remedy the units problem, it does nothing to ensure that we do not overfit correlated fMRI data.

We should also mention that while placement of λ can ensure a quadratic cost, there are several additional parameters involved in the generation of β. These additional parameters are present regardless of whether one uses a fully nonlinear model for β or a linear time-invariant one. One would like to estimate these parameters or integrate them out as Bayesian nuisance parameters (Sivia and Skilling, 2006), but this would come at a significant increase in computational cost. We continue to work on variants of the MRJI that address these difficulties, but for the type of rapid event-related experiment considered here we do not recommend its use.

We therefore recommend the PowR method. It requires iteration, but the all-zero starting guess gives a very efficient estimation with scaled conjugate gradient. The cost function is quartic rather than quadratic, but this is a relatively mild nonlinearity. Fig. 5 and the supplemental figures show that PowR yielded solution power that is well correlated with fMRI activity, particularly when considering the most active voxels, but which is not absolutely coincident with fMRI measurements. In all four subjects the scatterplots of fMRI power versus solution power (middle column of Fig. 5 and the supplemental figures) show this ability to not overfit to the fMRI data when it is weak. Also, when considering the MRI slices (third column of Fig. 5 and the supplemental figures) PowR yielded activity maps in which all four subjects show strong (but not exclusive) occipital activity, as we expect from this visual evoked response experiment. We also suggest that PowR attempts to take advantage of the best features of each data modality; it fits the temporal dynamics of EEG and the spatial dynamics of fMRI to obtain one high-resolution map.

We mention here that the activity we are localizing is not all stimulus-related activity; our procedures for selecting components from the EEG and fMRI data would exclude all non-evoked activity, that is, activity which is not phase-locked to the stimulus. If the stimuli modulate oscillatory activity unlocked in phase to that same stimulus, the methodology we present here would exclude that activity. To attempt to reveal and retain this non-evoked activity, we could potentially average the time–frequency power maps of the components (anonymous reviewer, private communication), employ spectral clustering (Makeig et al., 2004), or directly apply ICA in the frequency domain (Anemüller et al., 2003).

In going beyond the present study, issues of experimental design and data resolution are key. Slow versus fast event-related designs may require different joint inversion methods and we are currently conducting further empirical studies to address this issue. Also important are issues of data resolution; we have used a simple three sphere head model in this study, but detailed anatomical head models using cortical and scalp surface extraction from structural

Fig. 7. Joint inverse solution using PowR for four subjects participating in the dual evoked visual response experiment. Each basis function along with its loadings in a representative slice are shown, as in Fig. 6. Small loadings have again been omitted for clarity. Each row, labeled S1–S4, shows a different subject, and the columns pair a basis function with the corresponding loadings in a representative slice, as in Fig. 6. The grouping is not arbitrary. Subject S1 had only two task-related components. His/her entire solution is shown in row S1, and for the other subjects, we show the two basis functions most correlated with S1's directly below (grouped with a dotted box). Continuing with S2, in which we found three basis functions, we then vertically group this component with the component in the remaining subjects most similar as again judged by correlation. Heavy black arrows in row S4 indicate stimulus onsets (first frame, second frame); these onsets are the same for all basis function panels but suppressed to avoid clutter.

MRIs exist (Spinelli et al., 2000; Ermer et al., 2001). How much descriptive power is gained by using an anatomically detailed head model, and how does this choice interact with the signal-to-noise ratio in the EEG data? The EEG electrodes are also highly spatially correlated, suggesting the use of a lower-density EEG montage would result in the same quality of inverse solution. In addition, the EEG data could be temporally downsampled and the fMRI data spatially smoothed and/or resliced at lower resolution. All of these choices affect both the size of the subsequent joint inverse problem (and hence solution efficiency) as well as the nature of the solution obtained.

We would also like to be able to combine data from multiple participants into a group-level map. However, we believe doing so with these multimodal maps is a subtle issue that requires care. The challenge is the following: suppose as in Fig. 7 S1 has two components and S4 has five components. What does it mean to do a group average? We would like a procedure that finds group-wise consistent components and then generates a group-wise spatial map for those components also. However, to do this, we need to be able to determine our confidence in the components from each subject. If a component in S4 but not S1 is highly significant—meaning robustly present despite noise in the data and choices made during processing—the group analysis has to treat such an outlier as important. However, if the unique component is essentially lost in the noise (insignificant) then it may be ignored in the group average.

One basic measure of parameter error bars (for example in the loadings α) is the covariance matrix of the fitted parameters, which is the same as the inverse of the Hessian matrix of the cost function at the minimum. A more careful analysis would sample from the posterior density, but both of these methods ignore all the steps leading up to joint inversion (Fig. 1), many of which are not easily expressible in parametric form, so they cannot be properly sampled over in this way. Even for single imaging modality data the effect of choices in the data processing pipeline is a complex and active area of study (Strother et al., 2004; Strother, 2006). We believe a cross-validation approach would be best for determining robust solution features, and we are actively working on this.

The next phase of cognitive neuroscience is to go beyond studying local brain regions and to begin to learn about the global, distributed networks in the brain underlying cognitive activity (Bullmore and Sporns, 2009). In order to understand the dynamics within and between such networks, measurements with both high temporal and spatial resolution are essential. In addition, as researchers move towards more complicated, naturalistic stimuli (Malinen et al., 2007) in an effort to understand complex decision making, ever more sophisticated techniques for extracting relevant information from neuroimaging datasets will be required. Pushing EEG/fMRI fusion techniques to the limit will optimize our ability to extract information from these noninvasive, proxy measurements of the neuronal basis of cognition.

Acknowledgments

K.S.B. thanks T. Brookings and M. Fewings for useful discussions, A. Johnson for helping in the preparation of Fig. 7, and S. Lentz for providing MATLAB code for the filter in Eq. 5. The authors thank two anonymous reviewers for comments which helped us improve this manuscript. This work was supported by the David and Lucile Packard Foundation and the Institute for Collaborative Biotechnologies through contract no. W911NF-09-D-0001 from the U.S. Army Research Office.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2009.10.011.

References

Alliney, S., Ruzinsky, S., 1994. An algorithm for the minimization of mixed £1 and £2 norms with application to Bayesian estimation. IEEE Trans. Signal Process. 42 (3), 618–627.
Anemüller, J., Sejnowski, T.J., Makeig, S., 2003. Complex independent component analysis of frequency-domain electroencephalographic data. Neural Netw. 16, 1311–1323.
Berg, P., Scherg, M., 1994. A fast method for forward computation of multiple-shell spherical head models. Electroencephalogr. Clin. Neurophysiol. 90 (1), 58–64.
Brookings, T., Ortigue, S., Grafton, S., Carlson, J., 2009. Using ICA and realistic BOLD models to obtain joint EEG/fMRI solutions to the problem of source localization. Neuroimage 44, 411–420.
Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–198.
Buxton, R., Wong, E., Frank, L., 1998. Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. Magn. Reson. Med. 39, 855–864.
Calhoun, V., Adali, T., Pearlson, G., Kiehl, K., 2006. Neuronal chronometry of target detection: fusion of hemodynamic and event-related potential data. Neuroimage 30, 544–553.
Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26, 55–67.
de Peralta Menendez, R.G., Andino, S.L.G., Morand, S., Michel, C.M., Landis, T., 2000. Imaging the electrical activity of the brain : ELECTRA. Hum. Brain Mapp. 9, 1–12.
Eichele, T., Calhoun, V., Moosmann, M., Specht, K., 2008. Unmixing concurrent EEG-fMRI with parallel independent component analysis. Int. J. Psychophysiol. 67, 222–234.
Eichele, T., Calhoun, V.D., Debener, S., 2009. Mining EEG–fMRI using independent component analysis. Int. J. Psychophysiol. 73 (1), 43–61.
Ermer, J.J., Mosher, J.C., Baillet, S., Leahy, R.M., 2001. Rapidly recomputable EEG forward models for realistic head shapes. Phys. Med. Biol. 46, 1265–1281.
Flagg, C. N., Vermersch, J. A., Beardsley, R. C., 1976. 1974 MIT New England Shelf Dynamics Experiment (March, 1974) Data Report, Part II: The moored array. Tech. Rep. 76-1, Massachusetts Institute of Technology.
Grech, R., Cassar, T., Muscat, J., Camilleri, K.P., Fabri, S.G., Zervakis, M., Xanthopoulos, P., Sakkalis, V., Vanrumste, B., 2008. Review on solving the inverse problem in EEG source analysis. J. NeuroEng. Rehabil. 5 (1), 25.
Hallez, H., Vanrumste, B., Grech, R., Muscat, J., Clerq, W.D., Vergult, A., D ' Asseler, Y., Camilleri, K.P., Fabri, S.G., Huffel, S.V., Lemahieu, I., 2007. Review on solving the forward problem in EEG source analysis. J. NeuroEng. Rehabil. 4 (46), 29.
Hämäläinen, M., Ilmoniemi, R., 1994. Interpreting magnetic fields of the brain: minimum norm estimates. Med. Biol. Eng. Comput. 32, 35–42.
Hauk, O., 2004. Keep it simple: a case for using classical minimum norm estimation in the analysis of EEG and MEG data. Neuroimage 21, 1612–1621.
Huang, M.-X., Dale, A.M., Song, T., Halgren, E., Harrington, D.L., Podgorny, I., Canive, J.M., Lewis, S., Lee, R.R., 2006. Vector-based spatial-temporal minimum L1-norm solution for MEG. Neuroimage 31, 1025–1037.
Hyvarinen, A., Oja, E., 1997. A fast fixed-point algorithm for independent component analysis. Neural Comput. 9 (7), 1483–1492.
Hyvarinen, A., Oja, E., 2000. Independent component analysis: algorithms and applications. Neural Netw. 13, 411–430.
Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M.J., Iragui, V., Sejnowski, T. J., 2000. 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., 2001. Analysis and visualization of single-trial event-related potentials. Hum. Brain Mapp. 14, 166–185.
Jutten, C., Herault, J., 1991. Blind separation of sources, part 1: an adaptive algorithm based on neuromimetic architecture. Signal Process. 24, 1–10.
Kohler, T., Wagner, M., Fuchs, M., Wischmann, H., Drenckhahn, R., Thei en, A., 1996. Depth normalization in MEG/EEG current density imaging. 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 812–813.
LeVan, P., Urrestarazu, E., Gotman, J., 2006. A system for automatic artifact removal in ictal scalp EEG based on independent component analysis and Bayesian classification. Clin. Neuro-physiol. 117, 912–927.
Limeburner, R., Alessi, C. A., Beardsley, R. C., Rosenfeld, L. K., Lentz, S. J., Send, U., Winant, C. D., Allen, J. S. Jr., Brown, G. R. H., Irish, J. D., 1983. CODE-2 : moored array and large-scale data report. Tech. Rep. 85-35, Woods Hole Oceanographic Institution.
Liu, Z., He, B., 2008. fMRI–EEG integrated cortical source imaging by use of time-variant spatial constraints. Neuroimage 39, 1198–1214.
Liu, A.K., Belliveau, J.W., Dale, A.M., 1998. Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. Proc. Natl. Acad. Sci. U. S. A. 95, 8945–8950.
Makeig, S., Bell, A., Jung, T., Sejnowski, T., 1996. Independent component analysis of electroencephalographic data. Advances in Neural Information Processing Systems, 8, p. 7.
Makeig, S., Delorme, A., Westerfield, M., Jung, T.-P., Townsend, J., Courchesne, E., Sejnowski, T.J., 2004. Electroencephalographic brain dynamics following manually responded visual targets. PLoS Biol. 2 (6), 0747–0762.
Malinen, S., Hlushchuk, Y., Hari, R., 2007. Towards natural stimulation in fMRI—issues of data analysis. Neuroimage 35, 131–139.
McKeown, M.J., Makeig, S., Brown, G.G., Jung, T.-P., Kindermann, S.S., Bell, A.J., Sejnowski, T.J., 1998. Analysis of fMRI data by blind separation into independent spatial components. Hum. Brain Mapp. 6, 160–188.

Møller, M.F., 1993. A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw. 6, 525–533.
Moosmann, M., Eichele, T., Nordby, H., Hugdahl, K., 2008. Joint independent component analysis for simultaneous EEG–fMRI: principle and simulation. Int. J. Psychophysiol. 67, 212–221.
Oldfield, R., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113.
Ortigue, S., Thompson, J.C., Parasuraman, R., Grafton, S.T., 2009. Spatio-temporal dynamics of human intention understanding in temporo-parietal cortex: a combined EEG/fMRI repetition suppression paradigm. PLoS One 4, e6962.
Ou, W., Hamaläinen, M.S., Golland, P., 2009. A distributed spatio-temporal EEG/MEG inverse solver. Neuroimage 44, 932–946.
Pascual-Marqui, R., Michel, C., Lehmann, D., 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18 (1), 49–65.
Phillips, C., Rugg, M., Friston, K., 2002. Anatomically informed basis functions for EEG source localization: combining functional and anatomical. Neuroimage 16, 678–695.
Phillips, Mattout, J., Rugg, M., Maquet, P., Friston, K., 2005. An empirical Bayesian solution to the source reconstruction problem in EEG. Neuroimage 24, 997–1011.
Robinson, P.A., Drysdale, P.M., der Merwe, H.V., Kyriakou, E., Rigozi, M.K., Gemanoska, B., Rennie, C.J., 2006. Bold responses to stimuli: dependence on frequency, stimulus form, amplitude, and repetition rate. Neuroimage 31, 585–599.
Sekihara, K., Nagarajan, S., Poeppel, D., Miyashita, Y., 2001. Reconstructing spatio-temporal activities of neural sources using an MEG vector beamformer technique. IEEE Trans. Biomed. Eng. 48 (7), 760–771.
Sivia, D.S., Skilling, J., 2006. Data analysis: a bayesian tutorial. Oxford University Press.
Spinelli, L., Andino, S.G., Lantz, G., Seeck, M., Michel, C.M., 2000. Electromagnetic inverse solutions in anatomically constrained spherical head models. Brain Topogr. 13 (2), 115–125.
Stone, J.V., 2004. Independent component analysis: a tutorial introduction. MIT Press.
Strother, S., 2006. Evaluating fMRI preprocessing pipelines. IEEE Eng. Med. Biol. Mag. 27–41 March/April.
Strother, S., Conte, S.L., Hansen, L.K., Anderson, J., Zhang, J., Pulapura, S., Rottenberg, D., 2004. Optimizing the fMRI data-processing pipeline using prediction and reproducibility performance metrics. I. A preliminary group analysis. Neuroimage 23, S197–S207.
Ting, K.H., Fung, P.C.W., Chang, C.Q., Chan, F.H.Y., 2006. Automatic correction of artifact from single-trial event-related potentials by blind source separation using second order statistics only. Med. Eng. Phys. 28, 780–794.
Wunsch, C., 1996. The Ocean Circulation Inverse Problem. Cambridge University Press.