# Sparsity enables estimation of both subcortical and cortical activity from MEG and EEG

**Pavitra Krishnaswamy[a,b,c], Gabriel Obregon-Henao[d], Jyrki Ahveninen[a,e], Sheraz Khan[a,e,f], Behtash Babadi[g], Juan Eugenio Iglesias[a], Matti S. Hämäläinen[a,e,h,i,1,2], and Patrick L. Purdon[d,e,1,2]**

[a]Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA 02129; [b]Harvard–Massachusetts Institute of Technology Division of Health Sciences and Technology, Cambridge, MA 02139; [c]Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore 138632, Singapore; [d]Department of Anesthesia, Critical Care, and Pain Medicine, Massachusetts General Hospital, Boston, MA 02114; [e]Harvard Medical School, Boston, MA 02115; [f]Department of Neurology, Massachusetts General Hospital, Charlestown, MA 02129; [g]Department of Electrical & Computer Engineering, University of Maryland, College Park, MD 20742; [h]Department of Neuroscience and Biomedical Engineering, Aalto University School of Science, Espoo 02150, Finland; and [i]The Swedish National Facility for Magnetoencephalography (NatMEG), Department of Clinical Neuroscience, Karolinska Institute, Stockholm 17177, Sweden

Subcortical structures play a critical role in brain function. However, options for assessing electrophysiological activity in these structures are limited. Electromagnetic fields generated by neuronal activity in subcortical structures can be recorded noninvasively, using magnetoencephalography (MEG) and electroencephalography (EEG). However, these subcortical signals are much weaker than those generated by cortical activity. In addition, we show here that it is difficult to resolve subcortical sources because distributed cortical activity can explain the MEG and EEG patterns generated by deep sources. We then demonstrate that if the cortical activity is spatially sparse, both cortical and subcortical sources can be resolved with M/EEG. Building on this insight, we develop a hierarchical sparse inverse solution for M/EEG. We assess the performance of this algorithm on realistic simulations and auditory evoked response data, and show that thalamic and brainstem sources can be correctly estimated in the presence of cortical activity. Our work provides alternative perspectives and tools for characterizing electrophysiological activity in subcortical structures in the human brain.

MEG | EEG | subcortical structures | source localization | sparsity

**D**eep brain structures play important roles in brain function. For example, brainstem and thalamic relay nuclei have a central role in sensory processing (1, 2). Thalamocortical and hippocampal oscillations govern states of sleep, arousal, and anesthesia (3–8). Striatal regions are crucial for movement planning, while limbic structures like the hippocampus and amygdala drive memory, emotion, and learning (9–13). Altered signaling within the thalamus, striatum, hippocampus, and amygdala underlies pathologies such as autism, dementia, and depression (14). Much of our understanding of subcortical function comes from lesion studies and invasive electrophysiological recordings in animal models. Improved tools to characterize subcortical activity in humans would make it possible to analyze interactions between subcortical structures and other brain areas and could be used to better understand how subcortical activity relates to perception, cognition, behavior, and associated disorders.

At present, techniques for characterizing deep brain dynamics are limited. Invasive electrophysiological recordings (15) in humans are generally limited to regions that need to be monitored for clinical purposes. Functional magnetic resonance imaging (fMRI) can noninvasively measure deep brain activity with a spatial resolution of up to $\sim$1−2 mm, but cannot record fast signals or oscillations. Magnetoencephalography (MEG) and electroencephalography (EEG) noninvasively measure fields generated by neural currents with millisecond-scale temporal resolution (16), and M/EEG source estimation (17) is widely used to spatially resolve neural dynamics to within $\sim$0.5−2 cm in the cerebral cortex (18–21). However, it remains an open question whether M/EEG can be used to estimate neural currents in deep brain structures.

The anatomy of deep brain structures poses significant challenges for source estimation with M/EEG. Subcortical structures produce smaller scalp M/EEG signals because subcortical structures are farther from the scalp than cortical structures. In addition, neurons in subcortical structures can have a closed-field geometry that further weakens the fields observed at a distance (22). A second, perhaps more fundamental problem is that subcortical structures are enclosed by the cortical mantle. Thus, measurements arising from activity within deep brain structures can potentially be explained by a surrogate distribution of currents on the cortical surface. This implies that it would be very difficult to estimate subcortical activity when cortical activity is occurring simultaneously.

Clearly, if the scalp M/EEG signals generated by subcortical structures are too small to measure, it is not possible to estimate subcortical sources. But in many cases, subcortical sources do generate measurable scalp signals, and in these cases it might be possible to estimate those sources, if only their fields could be distinguished from those generated by cortical sources. In many neuroscience studies, salient cortical activity can be restricted to a set of well-circumscribed areas (23, 24). We introduce the hypothesis that if the cortical generators are sparse in this sense,

## Significance

Subcortical structures play a critical role in brain functions such as sensory perception, memory, emotion, and consciousness. There are limited options for assessing neuronal dynamics within subcortical structures in humans. Magnetoencephalography and electroencephalography can measure electromagnetic fields generated by subcortical activity. But localizing the sources of these fields is very difficult, because the fields generated by subcortical structures are small and cannot be distinguished from distributed cortical activity. We show that cortical and subcortical fields can be distinguished if the cortical sources are sparse. We then describe an algorithm that uses sparsity in a hierarchical fashion to jointly localize cortical and subcortical sources. Our work offers alternative perspectives and tools for assessing subcortical brain dynamics in humans.

www.manaraa.com

it might be possible to uniquely recover both the cortical and sub-cortical sources.

In this paper, we analyze M/EEG field patterns generated by cortical and subcortical sources and assess the extent to which sparse cortical and subcortical sources can be distinguished. We then introduce a hierarchical sparse estimation algorithm to illustrate how we can capitalize on these concepts to character-ize both cortical and subcortical activity. We demonstrate the algorithm's performance on simulated and experimental M/EEG data containing both cortical and subcortical activity.

## Theory

**Neural Sources and M/EEG Fields.** Primary neural currents (25), usually modeled with an ensemble of current dipoles, generate M/EEG fields **Y**,

$$\mathbf{Y}_{N \times T} = \mathbf{G}_{N \times M} \mathbf{X}_{M \times T} + \mathbf{V}_{N \times T}, \qquad [1]$$

where **G** is the gain matrix determined by the quasistatic approx-imation of Maxwell's equations, **X** contains the amplitudes of the current dipole sources, **V** is the noise, $N$ is the number of sen-sors, $M$ is the number of sources, and $T$ is the number of time points measured. To simplify notation, we assume that the data, the gain matrix, and the noise in Eq. 1 have been whitened to account for the spatial covariance $\mathbf{Q}_{N \times N}$ of the actual observa-tion noise (26), so that **V** is Gaussian with zero mean and identity covariance matrix $\mathbf{I}_{N \times N}$. When $T = 1$, we use notations **y** and **x** in lieu of **Y** and **X**.

We use high-resolution structural MRIs from healthy vol-unteers to delineate cortical surfaces and subcortical anatomic regions to define the locations and orientations of the elemen-tary dipole sources (*Materials and Methods*). We place sources on cortical surfaces and in subcortical volumes and cluster proxi-mal groups of dipoles into surface patches or volume subdivisions sized to homogenize signal strengths (*Materials and Methods*; Fig. 1 *A* and *B*; and *SI Appendix*, section SI I and Fig. S1). The result-ing set $\mathcal{B}$ of $K$ patches and subdivisions, together called divisions, constitutes the distributed brain source space.

We can then group the columns of **G** and rows of **X** according to these divisions and rewrite Eq. 1 as

$$\mathbf{Y}_{N \times T} = \sum_{k=1}^{K} \mathbf{G}_k \mathbf{X}_k + \mathbf{V}_{N \times T}, \qquad [2]$$

where $\mathbf{G}_k$ and $\mathbf{X}_k$ denote the gain matrix and source currents within the $k$th division, respectively. We compute $\mathbf{G}_k$ for each division $k$ in $\mathcal{B}$ and decompose it into a set of eigenmodes (*Mate-rials and Methods*). We denote gain matrices and source currents for a set of divisions $\mathcal{F} \subset \mathcal{B}$ by $\mathbf{G}_{\mathcal{F}} = \{\mathbf{G}_k\}$ and $\mathbf{X}_{\mathcal{F}} = \{\mathbf{X}_k\} \, \forall k \in \mathcal{F}$, respectively. Fig. 1 *C* and *D* illustrates field patterns for one cortical and one subcortical division.

**Fields Generated by Subcortical Sources Can Be Explained by Currents on the Cortex.** To analyze distinctions between subcortical and cortical fields, we investigated the extent to which MEG field patterns arising from subcortical currents can be explained by cortical surrogates.

We simulated field pattern $\mathbf{y}_{\mathcal{V}}$ corresponding to unit current in the VPL thalamus (Fig. 2 *A* and *B*) and assessed whether $\mathbf{y}_{\mathcal{V}}$ could be explained by some distribution of cortical source currents. Specifically, we fitted the subcortical field pattern with cortical sources; i.e., we computed the cortical minimum-norm estimate to explain $\mathbf{y}_{\mathcal{V}}$. We found that the resulting currents are small and broadly distributed across several cortical patches (Fig. 2*C*). Further, the goodness of fit between the field pattern for the cortical estimate (Fig. 2*D*) and $\mathbf{y}_{\mathcal{V}}$ was 98.5%, showing that $\mathbf{G}_{\mathcal{V}}$ can be explained by some combination of sources in the full dense cortical space.
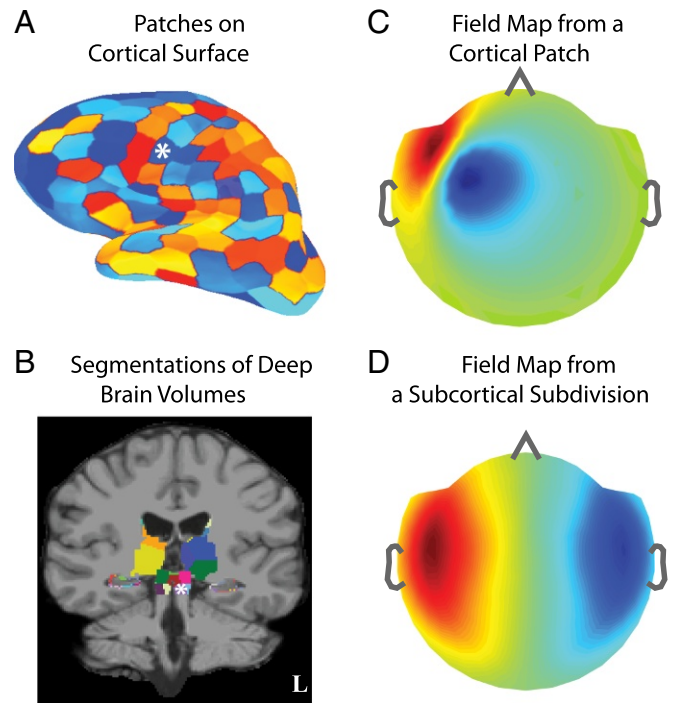


**Fig. 1.** Illustration of source spaces and field patterns. (*A*) Cortical source space $\mathcal{C}$ comprising patches (sized 650 mm$^2$) on the gray–white matter sur-face interface. (*B*) Subcortical source space $\mathcal{S}$ comprising volume subdivi-sions (sized $175-1,800$ mm$^3$) in the caudate, putamen, amygdala, thalamus, brainstem, and surface patches (sized 47 mm$^2$) on hippocampus. (*A* and *B*) The set of cortical and subcortical divisions $\mathcal{B} = \mathcal{C} \cup \mathcal{S}$ forms the full dis-tributed source space. (*C* and *D*) Example of noiseless MEG field pattern arising from activity in a frontotemporal cortical patch and a brainstem sub-division, respectively (white asterisks in *A* and *B*). Inflated surfaces and field maps have a left–right convention opposite to the MRI views. (*C* vs. *D*) Cor-tical fields tend to have more focal spatial patterns, while subcortical fields tend to be more distributed.

**Analysis of Forward Solutions.** To generalize the above result, we used principal angles (27, 28) to characterize the relation-ship between the cortical and subcortical field patterns. Principal angles quantify the correlation between linear subspaces, in this case the space spanned by MEG fields arising from sources in dif-ferent cortical and subcortical regions. For the example in Fig. 2, the maximum principal angle between subspaces spanned by the subcortical and cortical gain matrices was $0°$. Further, the maxi-mum principal angle between any subcortical gain matrix and the cortical gain matrix was $0°$. We conclude that the presence of the full cortical source space makes it impossible to unambiguously estimate currents in simultaneously active subcortical sources.

**Sparsity Makes It Possible to Distinguish Fields from Subcortical and Cortical Sources.** We next studied the extent of subspace corre-lation when subcortical sources are active together with only a small subset of cortex. As an example, we examined the scenario of median-nerve somatosensory stimulation, which elicits activity in ventroposterolateral (VPL) thalamus, primary and secondary sensory cortices (S1, S2), and posterior parietal cortex (PPC) (29), and analyzed forward solutions for a source space encom-passing these brain areas (Fig. 3*A*).

Specifically, we considered all possible configurations of sub-cortical and cortical currents in these divisions, computed prin-cipal angles between the subcortical and cortical gain matri-ces corresponding to each possible configuration (*Materials and Methods*), and plotted the distribution of angles (Fig. 3*B*). A large proportion of the principal angles are high (median $43.9°$),
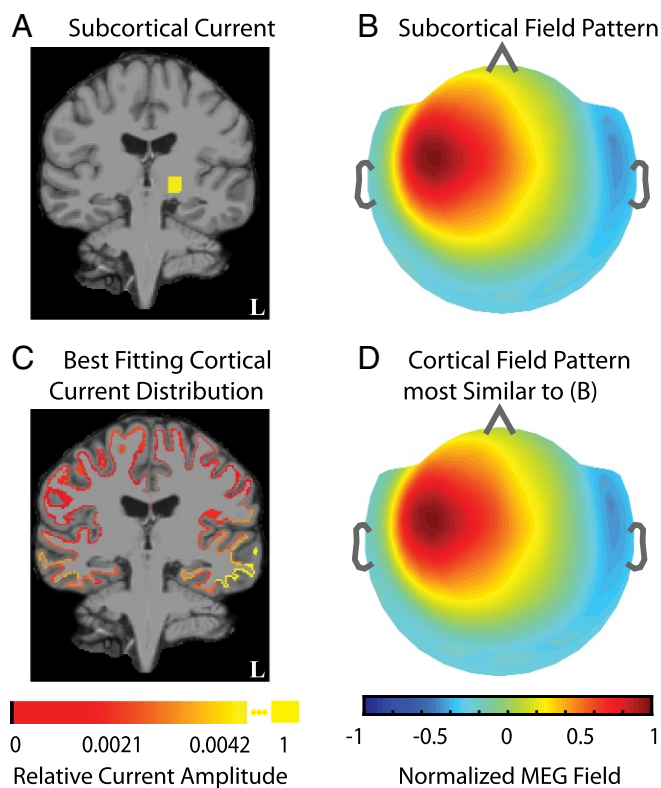
Krishnaswamy et al.

www.manaraa.com

**Fig. 2.** Fields generated by subcortical sources can be explained by currents on the cortex. (*A* and *B*) Example of unit source current in left ventroposterolateral (VPL) thalamus (sized 1.5 cm³) and the corresponding noiseless MEG field pattern. (*C* and *D*) Distribution of currents on cortical surface patches (sized 650 mm²) that reproduce the MEG field pattern generated by the subcortical source. *C* and *D* show the fitted cortical currents and MEG field pattern. The source current plots in *A* vs. *C* are the resultant currents from dipoles within a subdivision. The field maps in *B* and *D* are normalized. The fitted cortical field pattern is indistinguishable from the simulated subcortical pattern. This analysis illustrates how a subcortically generated field can be explained by some distribution of cortical currents.

indicating that many different configurations of activity within the sparse cortical and subcortical divisions can be distinguished from one another. We also computed principal angles for all mutually exclusive configurations of activity within the cortical divisions in Fig. 3*A* and found comparable principal angles (median 63.5°). Therefore, in principle, the problem of distinguishing subcortical sources from sparse cortical sources is similar in difficulty to that of distinguishing sparse cortical sources from one another. We also illustrate typical subcortical and cortical field patterns (*Materials and Methods*) for source current configurations corresponding to the various angles in this distribution (Fig. 3 *C* and *D*). Subcortical and cortical field patterns with principal angles as low as 45° are clearly distinguishable.

This example represents a conservative scenario, but illustrates an approach for characterizing the extent to which subcortical sources can be resolved for any given cortical source distribution. We also found similar trends for other more general cases (*SI Appendix*, section SI II and Figs. S2 and S3). These results lead us to conclude that spatial sparsity constraints can enable distinctions between cortical and subcortical field patterns. Based on this analysis, we introduce and test an inverse algorithm that uses a sparse cortical representation to achieve localization of simultaneous subcortical and cortical activity.

## Inverse Algorithm

**Electromagnetic Inverse Problem.** The electromagnetic inverse problem is to estimate source currents $\mathbf{X}$ underlying M/EEG data $\mathbf{Y}$, given forward gain matrix $\mathbf{G}$ for divisions distributed across the brain. This inverse problem is commonly solved using the linear $l_2$ minimum-norm estimator (MNE),

$$\hat{\mathbf{X}}(\mathbf{G}, \mathbf{Y}) = \mathbf{W}^{\text{MNE}}(\mathbf{G})\mathbf{Y}, \qquad [3]$$

where $\hat{\mathbf{X}}(\mathbf{G}, \mathbf{Y})$ is an estimate of $\mathbf{X}$, and the MNE estimator $\mathbf{W}^{\text{MNE}}$ is a function of $\mathbf{G}$. The performance of $\mathbf{W}^{\text{MNE}}$ can be assessed using the resolution matrix

$$\mathbf{K}(\mathbf{G}) = \mathbf{W}^{\text{MNE}}(\mathbf{G})\mathbf{G}. \qquad [4]$$

The ideal $\mathbf{K} = \mathbf{I}$ corresponds to a $\mathbf{W}^{\text{MNE}}$ which exactly recovers the current locations and amplitudes, in the absence of noise (30). In practice, source estimates are biased and more extended than the true sources. This nonideal behavior can be analyzed using the spatial dispersion (SD) and dipole localization error (DLE) metrics (31) (*Materials and Methods*), which indicate how far the inverse solution for a given source spreads from the actual source location. We analyze the performance of the MNE on the subcortical and cortical source estimation problems and describe our proposed inverse algorithm.

**Estimation Performance—Distributed Cortical and Subcortical Sources.** We first considered the problem of estimating neural currents in a set of divisions $\mathcal{B}$ distributed across both cortical and subcortical structures in the brain (Fig. 4*A*). Fig. 4*B* shows the MNE resolution matrix $\mathbf{K}^{\text{MNE}}(\mathbf{G}_{\mathcal{B}})$. We found that estimates for the cortical sources concentrate around the diagonal (Fig. 4*B*, upper left), implying good resolution for cortical sources. On the other hand, estimates for subcortical sources have low amplitudes on the diagonal and instead spread to cortical sources (Fig. 4*B*, upper right and lower left). Fig. 4*C* shows the distribution of SD and DLE across all divisions. The median SD is 4.23 cm (close to the radius of the human brain), and the median DLE is 1.89 cm. The cortical DLE is ∼0.5 cm and the subcortical DLE is in the 2- to 3-cm range. These findings are consistent with our principal angle analyses (Fig. 2).

**Estimation Performance—Sparse Cortical and Distributed Subcortical Sources.** Earlier we found the principal angles improve when only sparse subsets of cortical sources are active alongside deep sources (Fig. 3). Therefore, we assessed whether the resolution of the MNE inverse solution improves similarly. We considered the somatosensory stimulation example from Fig. 3 and constructed a composite source space $\mathcal{B}_r$ comprising the sparse cortical divisions in Fig. 3*A* alongside all subcortical divisions (Fig. 4*D*). Fig. 4*E* shows the MNE resolution matrix $\mathbf{K}^{\text{MNE}}(\mathbf{G}_{\mathcal{B}_r})$. We see that the estimates for subcortical sources do not spread significantly to the cortex (Fig. 4*E*, low values for upper right and lower left). However, the estimates for subcortical sources tend to spread across the subcortical source space (Fig. 4*E*, lower right, off-diagonal portions). Fig. 4*F* shows the resolution error metrics across all divisions in $\mathcal{B}_r$. The median SD is 2.10 cm, and the median DLE is 0.805 cm. This is an improvement from the previous case, but still not as accurate as needed to resolve many subcortical sources. Given a sparse cortical source space as a starting point, we anticipate that it might be possible to use a subsequent sparse estimation step to reduce spread among subcortical sources.

**Estimation Performance—Sparse Cortical and Sparse Subcortical Sources.** Sparse estimation procedures based on $l_1$-norm minimization, projection pursuit, and/or Bayesian theory are effective at pruning out spurious features, while identifying relevant
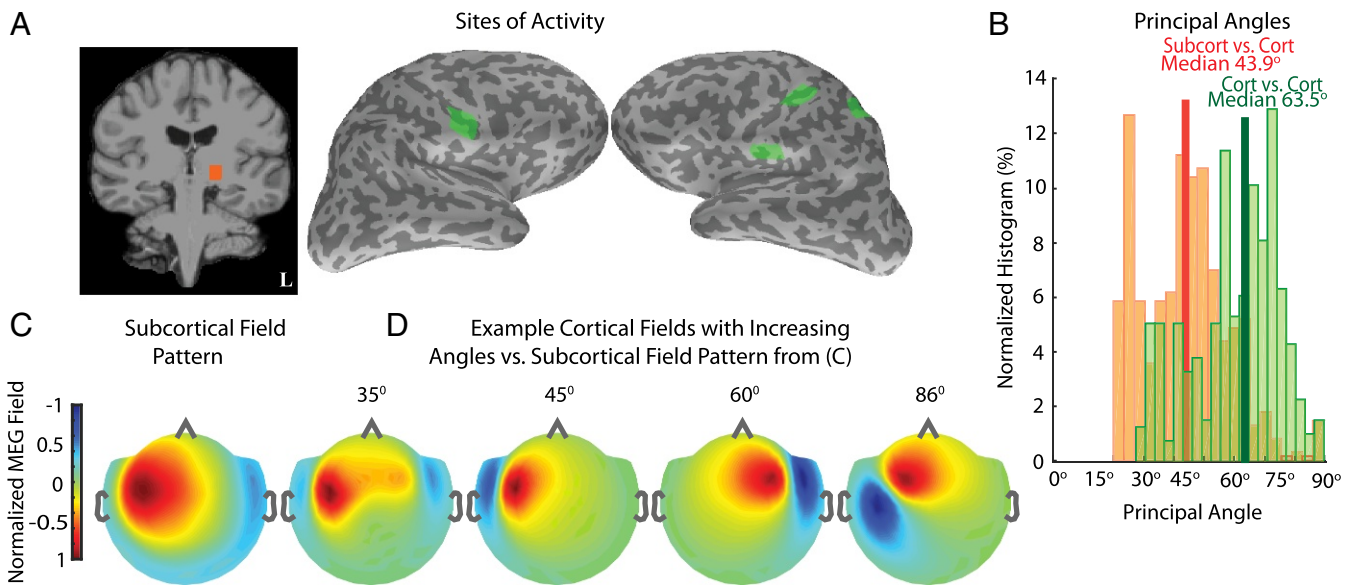
Krishnaswamy et al.

**Fig. 3.** Sparsity makes it possible to distinguish fields from subcortical and cortical sources. (*A*) Sources of activity derived from stimulation of the right median nerve: left VPL thalamus, primary and secondary somatosensory areas (S1, bilateral S2), and posterior parietal cortex (PPC). (*B*) Normalized histogram of principal angles, which quantify the correlation between fields arising from all representative combinations of activity within this neurophysiological source space. The orange histogram shows the distribution of subcortical vs. cortical angles, while the green histogram shows the distribution of cortico-cortical angles. (*C*) MEG field pattern resulting from activity in VPL thalamus. (*D*) Representative fields from example cortical source sets, whose gain matrices have the indicated principal angles with the subcortical gain matrix, that best fit the subcortical field pattern in C. All field map color scales are normalized to emphasize the spatial patterns. The spatial profiles of the cortical and subcortical MEG field patterns are distinct, even for a principal angle of 30°, and substantially so for principal angles >45°. These distinctions suggest the feasibility of resolving simultaneous subcortical and cortical activity.

sparse features in noisy high-dimensional problems (32–35). We have recently shown that subspace pursuit can accurately estimate multiple sparse cortical sources underlying MEG data (36). Thus, we assessed whether a similar subspace pursuit algorithm could reduce the spread among spurious subcortical sources and enable improved resolution for subcortical source estimates.

We continue with our analysis of the resolution matrix in the composite source space $\mathcal{B}_r$ (Fig. 4*G*, faded background), this time with subspace pursuit. Since the subspace pursuit algorithm is nonlinear, a closed-form resolution matrix in the sense of Eq. 4 does not exist; thus the performance of subspace pursuit must be characterized empirically instead. To this end, we simulated unit currents $\mathbf{x}_i$ in each $i$th division $\mathcal{B}_r(i)$ within $\mathcal{B}_r$, one at a time, to generate corresponding noiseless field patterns $\mathbf{y}_i = \mathbf{G}_{\mathcal{B}_r(i)}\mathbf{x}_i$. Then, for each $\mathbf{y}_i$, we performed subspace pursuit and constructed an empirical resolution matrix $\mathbf{K}^{SP}(\mathbf{G}_{\mathcal{B}_r})$ (*Materials and Methods*). The resulting matrix, shown in Fig. 4*H*, has a near-diagonal structure for the majority of cortical and subcortical sources. Moreover, Fig. 4*I* shows SD and DLE, across all divisions in $\mathcal{B}_r$. The median SD and DLE are the same and equal to 0.737 cm. This is a substantial improvement over previous solutions that do not use sparsity constraints (Fig. 4*H* vs. Fig. 4 *C* and *E*).

**Hierarchical Subspace-Pursuit Inverse Algorithm.** The above results suggest that it is possible to resolve both cortical and subcortical sources by applying sparsity in both domains. In previous work, we developed a sparse estimation algorithm for cortical divisions (36), in which sets of cortical divisions were nested in successively finer resolutions (i.e., smaller patches or divisions), and subspace pursuit was applied to derive sparse estimates in successively finer resolutions, which formed a hierarchy from coarse to fine resolution. We therefore intuited that subcortical sources could be viewed as an additional, ultimate step in this hierarchical refinement process, achieved by adding a set of subcortical divi-

sions to the final set of sparse cortical sources and applying subspace pursuit.

**Inverse Algorithm.** Inputs: Data $\mathbf{Y}$, distributed gain matrix $\mathbf{G}_\mathcal{B}$, and target sparsity level $L$.

Notation: Denote the distributed cortical and subcortical source spaces as $\mathcal{B}_C$ and $\mathcal{B}_S \subset \mathcal{B}$, respectively. Denote $\mathcal{H}_r \subset \mathcal{B}_r$ as the set of $L$ brain divisions whose estimated neural currents $\hat{\mathbf{X}}_{\mathcal{H}_r}$ best explain data $\mathbf{Y}$.

1. Do subspace pursuit on the distributed cortical source space $\mathcal{B}_C$: $[\mathcal{H}_C, \hat{\mathbf{X}}_{\mathcal{H}_C}] = SP(\mathbf{Y}, \mathbf{G}_{\mathcal{B}_C}, L)$.

2. Construct $\mathcal{B}_{C,\text{refined}} = \mathcal{H}_C \cup$ neighbors of $\mathcal{H}_C$ in a finer subdivision of cortical patches.

3. Repeat subspace pursuit on the coarse-to-fine hierarchy of cortical source spaces $\mathcal{B}_{C,\text{refined}}$: $[\mathcal{H}_{C_{sp}}, \hat{\mathbf{X}}_{\mathcal{H}_{C_{sp}}}] = SP(\mathbf{Y}, \mathbf{G}_{\mathcal{B}_{C,\text{refined}}}, L)$.

4. Construct the composite space of sparse cortical sources and distributed subcortical sources: $\mathcal{B}_r = [\mathcal{H}_{C_{sp}} \cup \mathcal{B}_S]$.

5. Repeat subspace pursuit on the composite sparse space $\mathcal{B}_r$: $[\mathcal{H}_r, \hat{\mathbf{X}}_{\mathcal{H}_r}] = SP(\mathbf{Y}, \mathbf{G}_{\mathcal{B}_r}, \alpha L)$, where $\alpha > 1$.

Outputs: Cortical and subcortical source locations $\mathcal{H} = \mathcal{H}_r \subset [1, 2, \dots, K]$; and the estimated time courses of neural currents at these locations $\hat{\mathbf{X}}_\mathcal{H} = \hat{\mathbf{X}}_{\mathcal{H}_r}$.

***Subspace pursuit (steps 1, 3, and 5).*** For a source space comprising a subset of brain divisions $\mathcal{F} \subset \mathcal{B}$ with gain matrices $\mathbf{G}_\mathcal{F}$, subspace pursuit (SP) estimates the locations and time courses of neural currents to explain data series $\mathbf{Y}_{N \times T}$:

$$[\mathcal{H}, \hat{\mathbf{X}}_\mathcal{H}] = SP(\mathbf{Y}, \mathbf{G}_\mathcal{F}, L). \qquad [5]$$

The pursuit procedure finds a sparse subset of dictionary $\mathbf{G}_\mathcal{F}$ that best explains data $\mathbf{Y}$, computes residuals remaining to be explained, and iterates to find new relatively uncorrelated subsets of $\mathbf{G}_\mathcal{F}$ that best explain these residuals, until all matching subsets of $\mathbf{G}_\mathcal{F}$ are found (34–36) (*SI Appendix*, section SI III).
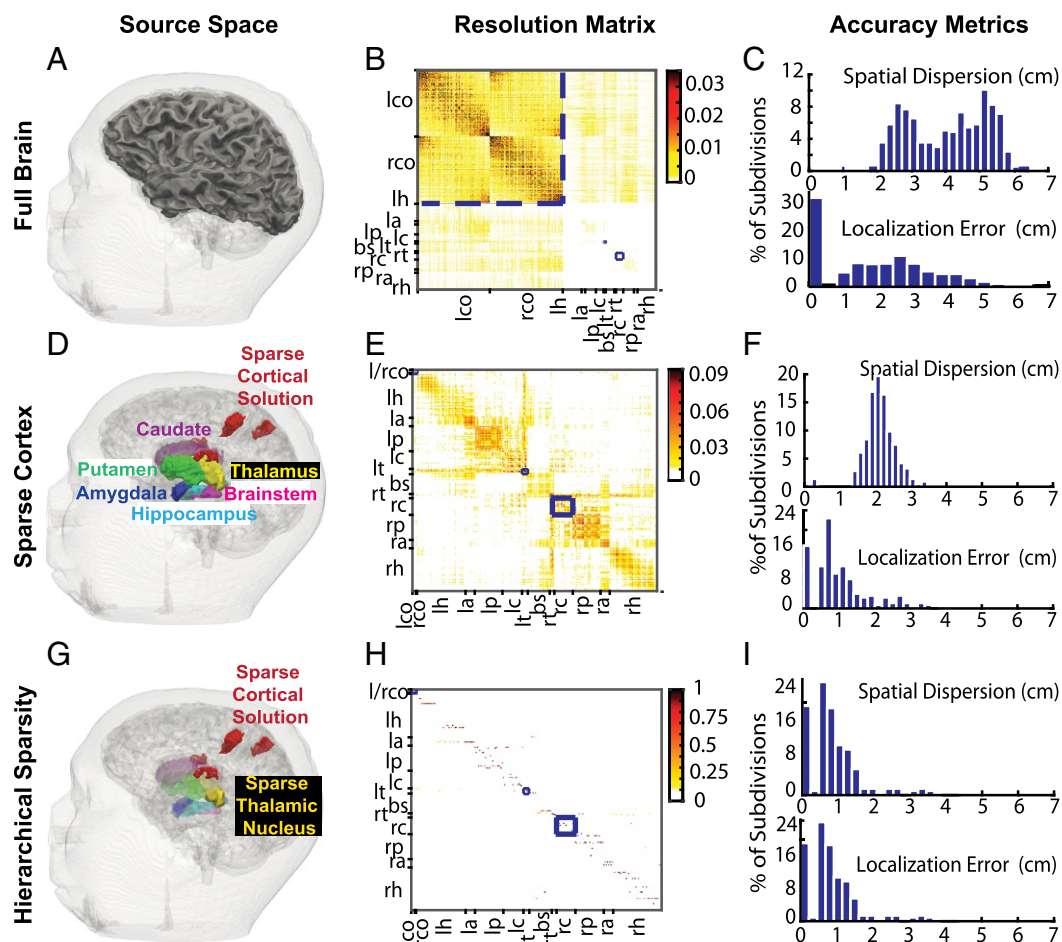
Krishnaswamy et al.

www.manaraa.com

**Fig. 4.** An analysis of how sparsity and hierarchy influence subcortical source estimation. (*A*) Illustration of all brain divisions considered. (*B*) Minimum-norm estimator (MNE) resolution matrix for the source space in *A*. (*C*) Summary dispersion and error metrics for the resolution matrix in *B*. Cortical estimates concentrate around the diagonal (low localization error), whereas subcortical estimates spread significantly to the cortex (high spatial dispersion). (*D*) A reduced space composed of sparse cortical regions that generate somatosensory evoked potentials combined with all subcortical volumes. (*E* and *F*) MNE resolution matrix and associated performance metrics for the reduced source space in *D*. The sparse subset of the cortical source space allows subcortical activity to be estimated, albeit with significant spread to nondiagonal regions. (*G*) Final sparse cortical and subcortical source regions identified using an inverse solution employing sparsity constraints. The faded subcortical regions show the hierarchically reduced subcortical source space, while the foreground subcortical regions show estimated sources in the thalamus. (*H* and *I*) Empirical resolution matrix (one active source per column) and associated performance metrics for the sparse solution. Estimates mostly concentrate on and around the diagonal for both cortical and subcortical sources. *B*, *E*, and *H* show left/right (l/r) cortex (l/rco), hippocampus (r/lh), amygdala (r/la), putamen (r/lp), caudate (r/lc), thalamus (r/lt), and brainstem (bs). All resolution matrices order sources based on physical proximity. Therefore, when sources are estimated accurately, the resolution matrix has a diagonal appearance. The blue boxes are used to delineate the position of the cortical, left thalamic, and right caudate sources in the resolution matrices. The changes in the color-scale range highlight the $3-10\times$ increase in recovered source amplitude when sparse estimation is applied across progressively refined hierarchies. Overall, hierarchical sparsity enables focal spatial resolution with minimal dispersion (or point spread) for inverse solutions incorporating both cortical and subcortical sources.

***Hierarchical construction (steps 2 and 4).*** We first apply SP on a distributed cortical source space to identify the subset of cortical divisions that best explain the measured fields. Then, to improve accuracy of this sparse subset of cortical sources, we perform SP across a coarse-to-fine hierarchy of cortical source spaces (36) (*SI Appendix*, Fig. S4). Subsequently, we construct a composite space of the sparse cortical source estimates and distributed subcortical sources and use SP on this composite sparse space to identify the subset of subcortical and cortical sources that best explain the data. This process of using sparse estimation across increasingly refined hierarchies enables systematic reduction of the distributed source space and the gain matrix: pruning out sources not important for explaining the data, implicitly decorrelating the columns of the gain matrix, and concentrating estimates into subsets of the brain whose neural currents best explain the data. Overall, given data **Y**, distributed gain matrix $\mathbf{G}_{\mathcal{B}}$, and target sparsity level $L$, the hierarchical SP algorithm identifies

the sparse subset $\mathcal{H}$ that specifies locations for both cortical and subcortical sources and estimates the time courses $\mathbf{X}_{\mathcal{H}}$ of neural currents at those locations.

## Data Examples

We illustrate the performance of the algorithm by analyzing noisy evoked response simulations and experimental data. First, we preprocess the measurements **Y** and estimate the noise covariance matrix **Q**. Next, we use the MRI data to construct the distributed source space, i.e., brain divisions $\mathcal{B}$, and compute the forward solutions $\mathbf{G}_{\mathcal{B}}$. Then, we use the hierarchical SP inverse algorithm to estimate the locations and time courses of sources that best explain the M/EEG data. We perform all data analyses at the individual subject level. We specify the target sparsity level $L$ for SP empirically, based on a conservative approximation of the expected number of active divisions. We maintain $L$ across the cortical hierarchies and increase it by a factor of $\alpha \sim 1.5-2$

Krishnaswamy et al.

for the final hierarchy, i.e., the composite hybrid source space $\mathcal{B}_r$ comprising sparse cortical sources and distributed subcortical sources. Estimates displayed in this section are obtained at the final hierarchy level $\mathcal{B}_r$. We note that $L$ specifies the number of eigenmodes included in the sparse solution and that currents from a given patch or subdivision may be represented by more than one eigenmode. As a result, the number of subdivisions seen on the topographical maps may be less than $L$.

**Somatosensory Evoked Simulations.** We simulated MEG evoked responses mimicking those elicited by electrical stimulation of the right median nerve at the wrist. Fig. 5 *A* and *B* illustrates the spatial and temporal features of simulated fields in the sensor space. The simulated fields include additive Gaussian noise, with a signal-to-noise ratio (SNR) of 7 dB, similar to resting eyes-open recordings. Fig. 5 *C* and *D* display the spatial and temporal patterns of simulated currents in the source space. Specifically, the evoked responses comprise early currents in the left somatosensory region of thalamus (L Som Th), followed by currents in the left primary somatosensory cortex (L S1, near the postcentral gyrus), and later currents in the left PPC (L PPC) and bilateral secondary somatosensory cortices (L/R S2). The simulated thalamic current time course has a periodic on/off pattern up to 250 ms poststimulus. This pattern was chosen to assess source estimation performance for the challenging case of phasic, temporally overlapping subcortical and cortical activity. We

note that the MEG fields arising from thalamic source currents (e.g., 0−15 ms in Fig. 5*B*) lie below the observation noise and are significantly smaller than those arising from cortical currents (e.g., 70−100 ms in Fig. 5*B*).

For source estimation, we used a source space different from that used for the simulation, to recreate a scenario closer to what might occur in practice, in which the true generating sources and the source space parcellation might not correspond. We set the sparsity level $L$ to 8 and 12 for the cortical and hybrid hierarchies, respectively. The procedure refines source current estimates across cortical hierarchies (*SI Appendix*, Fig. S5) and culminates in the final hybrid hierarchy $\mathcal{B}_r$. The final spatial distributions and time courses of estimated source currents (Fig. 5 *E* and *F*) closely resemble those of the simulated ground truth (Fig. 5 *C* and *D*) for both cortical and subcortical sources. The estimated left thalamic time course in Fig. 5*F* matches the simulation in shape and phase and further is not contaminated by leakage from cortical sources.

Further, we found that our algorithm offers significant gains in performance compared with other methods that do not use principles of sparsity and hierarchy (*SI Appendix*, Figs. S6 and S7). We conclude that both sparsity and hierarchy are necessary to accurately resolve locations and time courses of the thalamic source currents (*SI Appendix*, Fig. S6). Further, using sparsity in a hierarchical fashion helps recover the true distribution of mean source activity across anatomic regions (*SI Appendix*, Fig. S7).
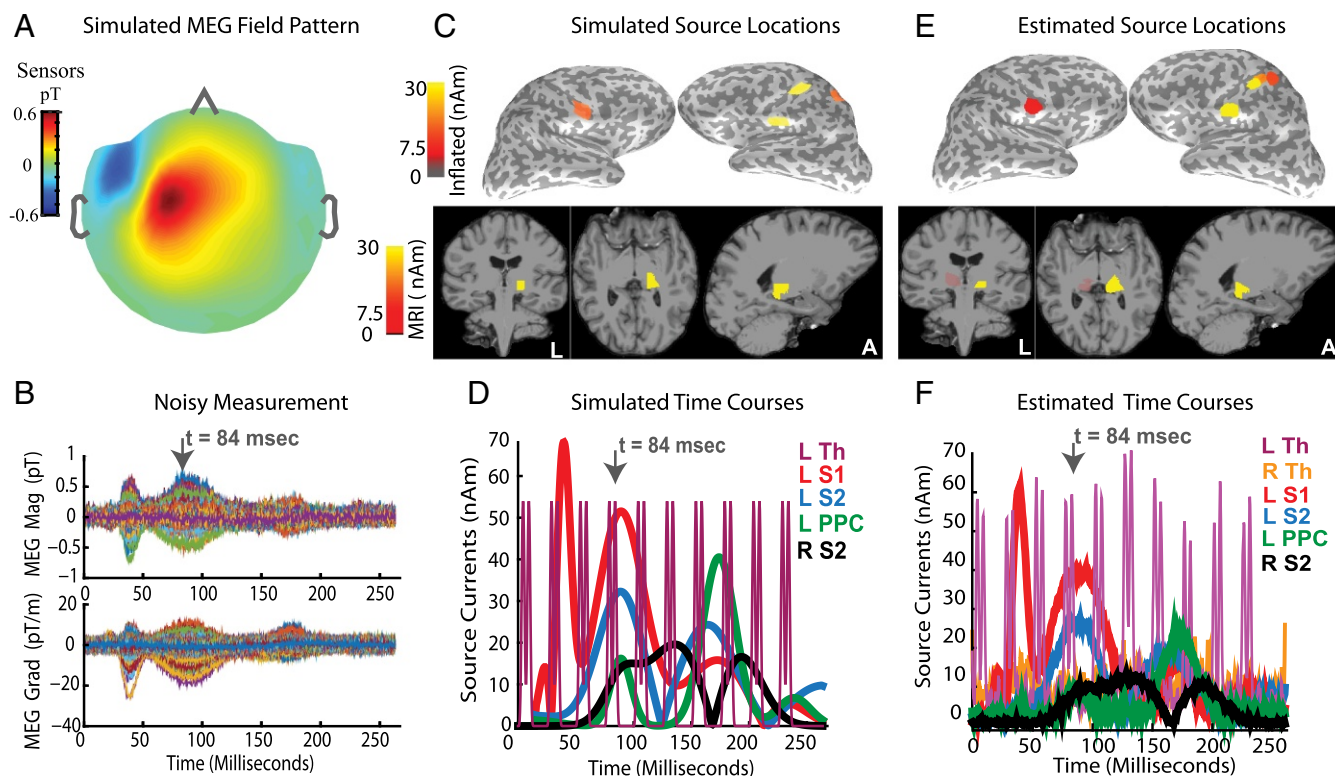


**Fig. 5.** Sparse hierarchical estimates recover simulated somatosensory responses. (*A* and *B*) Spatial distribution and time courses (one color per channel) of simulated MEG fields in sensor space. (*C* and *D*) Spatial distribution and time courses of simulated source currents in source space. Inflated views show sources located in somatosensory (S1, S2) and parietal (PPC) cortices. MRI views show thalamic source locations. The sagittal section passes through the left thalamus. The somatosensory thalamus (Th) is activated in a periodic on/off pattern. (*D* vs. *B*) While the cortical source currents contribute large-amplitude MEG signals, fields due to thalamic sources are not visible above the simulated noise. (*E* and *F*) Spatial distribution and time courses of estimated source currents in the source space. All topographical snapshots are at 84 ms (top gray arrows in time-course plots), the color scales and slice locations are the same in *C*–*E*, and all source currents are plotted in terms of the resultant magnitudes across dipoles within each region in the key. (*E* and *F* vs. *C* and *D*) Estimated source locations and time courses closely match the simulated ground truth. The thalamic source estimate follows the true phasic on/off pattern. Although there is a stray source estimate in right thalamus, it is weak and relatively constant over time. In *SI Appendix*, Figs. S6 and S7 compare the performance of our algorithm to alternatives that do not use sparsity and hierarchy.

**Auditory Evoked Response Experiments.** We recorded auditory evoked responses elicited by binaural stimulation with a train of clicks (37, 38) during the resting eyes-open condition simultaneously with MEG and EEG. Auditory responses comprise distinct M/EEG peaks at established latencies corresponding to a progression of activity from the cochlea, through the inferior colliculus in the brainstem, to the auditory cortex (37), and thus serve as a suitable test case for validating a subcortical source estimation algorithm.

The M/EEG evoked responses are shown in Fig. 6 *A* and *B*. We see early auditory brainstem response (ABR) peaks in both EEG and MEG at 5.8 ms and 6.2 ms, a low-amplitude Po feature in the EEG at ∼10 ms, and prominent midlatency response (MLR) peaks Na-Pa in MEG and EEG channels at 18−25 ms poststimulus. The ABR peaks are consistent with the brainstem wave V known to arise from the inferior colliculus (IC), the Po feature marks the end of brainstem components, and the Na-Pa peaks correspond to cortical responses known to arise in the auditory cortex (37).

We performed hierarchical SP and set sparsity levels $L$ to 4 and 8 for the cortical and hybrid hierarchies, respectively. Fig. 6 *C* and *E* shows localization of the 18- to 25-ms Na-Pa MLR peaks to bilateral auditory cortices and the associated time courses. The auditory areas compose the reduced cortical source space, which, along with the distributed subcortical sources, forms the hybrid source space $\mathcal{B}_r$ for estimation of deep sources underlying the ABR data. Fig. 6 *D* and *F* illustrates the localization of the wave V ABR peaks to bilateral ICs. Although the recorded wave V peaks do not have very high SNR, the source time course at the IC peaks at ∼5−6 ms and drops off after 10 ms, as expected. Consistent results in another subject are shown in (*SI Appendix*, Fig. S9). We compared performance to algorithms that do not use sparsity and hierarchy (*SI Appendix*, Fig. S10) and found that hierarchical SP is necessary to estimate specific subcortical sources even for filtered recordings containing temporally separated early latency responses.

## Discussion

The extent to which subcortical activity can be estimated from M/EEG measurements has been a controversial subject. Our key finding is that the M/EEG fields from subcortical sources can be distinguished from those generated by the cortex when the underlying cortical activity is sparse, i.e., limited to a subset of cortical regions. In this scenario, the problem of distinguishing
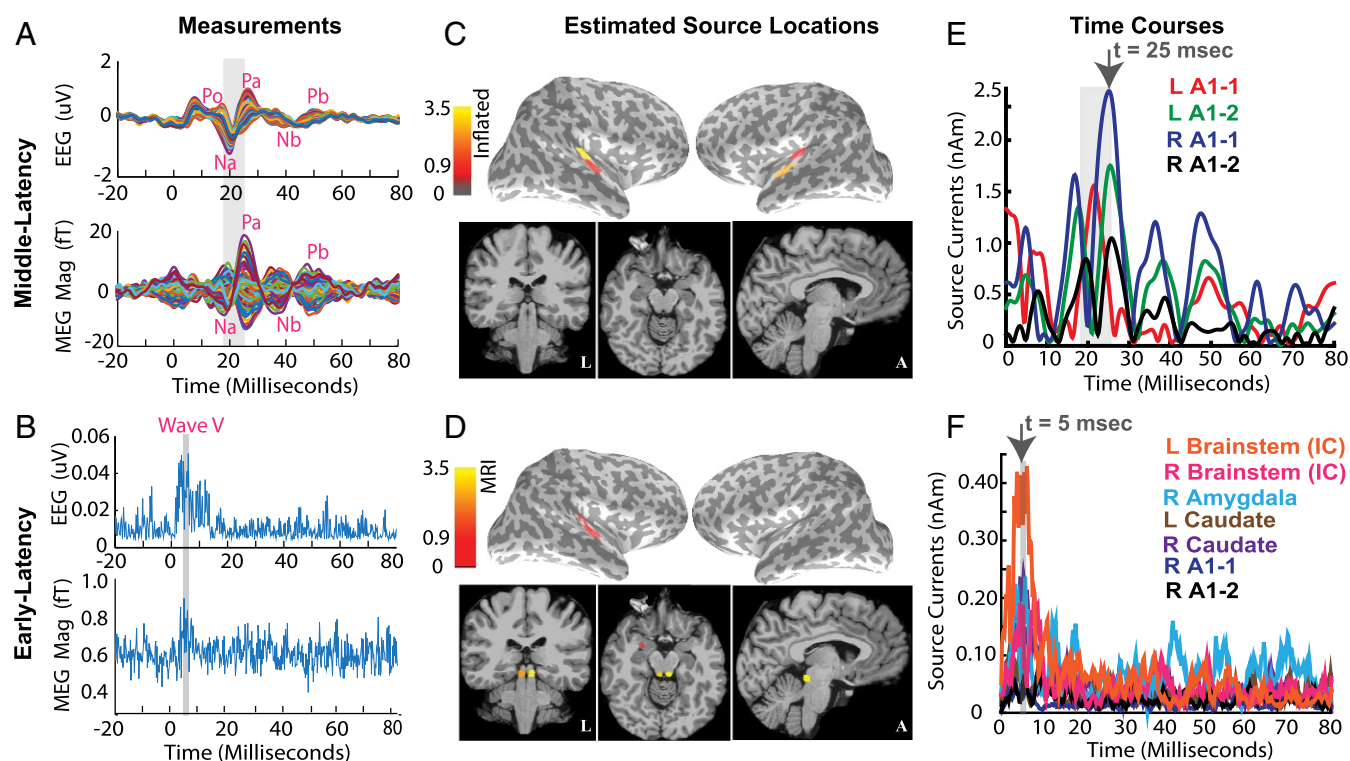


**Fig. 6.** Cortical and subcortical source estimates for evoked auditory responses. Stimulus-locked average auditory evoked responses were recorded from a healthy volunteer presented with a broadband click train stimulus. Time courses are averages across 11,170 epochs filtered between 500 Hz and 1,625 Hz for the auditory brainstem response (ABR), and between 30 Hz and 300 Hz for the middle latency response (MLR). (*A*) MLR time courses displayed across channels, one color per channel. Red labels denote common peaks occurring at the expected poststimulus latencies. The Na and Pa peaks (shaded gray section) are particularly prominent. (*B*) ABR time courses rectified and averaged across channels. The shaded gray section marks the 5.0- to 6.5-ms period poststimulus, when peaks consistent with ABR wave V appear in the recordings. (*C* and *E*) Sparse cortical estimates for middle-latency recordings (30−300 Hz): snapshots at 25 ms (top black arrow, *E*). The activity is localized to the Heschl's gyrus and superior temporal gyrus, consistent with auditory cortical processing. The source time courses from these areas have peaks consistent with the Na and Pa peaks in the scalp recordings in *A*. (*D* and *F*) Sparse hierarchical estimates for early-latency recordings (500−1,625 Hz), obtained using a source space comprised of sparse subsets of cortex in *C* and the distributed subcortical space. The spatial plots display the source activity at 5 ms (top black arrow, *F*). The activity is localized primarily to the inferior colliculus. A weak stray source is also seen in right amygdala. The brainstem source time courses show peaks consistent with the ABR wave V peaks in *B*. The color scales and slice locations are maintained for topographical snapshots across *C* and *D*, and all source currents are resultant magnitudes across dipoles within each region in the key. The color scales in *C* and *D* have units nAm and 0.1 nAm, respectively. Overall, our sparse hierarchical algorithm recovers cortical and subcortical sources consistent with the auditory stimuli presented. *SI Appendix*, Fig. S10 compares the performance of our algorithm to alternatives that do not use sparsity and hierarchy.

Krishnaswamy et al.

subcortical from cortical sources has a similar level of ambiguity to that of resolving different cortical sources. To demonstrate how this insight might be practically applied, we developed a sparse hierarchical algorithm to estimate both sparse cortical and subcortical sources.

It is known that deep and superficial sources exhibit different M/EEG field patterns (39), but the degree to which this information could be used to resolve multiple distributed subcortical and cortical sources has remained unclear. Analyses of cortical and subcortical field patterns assuming that entire structures can be simultaneously active have provided evidence for substantial correlation (40), consistent with our Fig. 2. However, we reasoned that it would be unlikely to observe synchronous activity, the major determinant of MEG/EEG (41), simultaneously within the entirety of cortex and any given subcortical structure. Thus, we analyzed sparse subdivisions of cortical and subcortical structures and found clear distinctions in the ensuing field patterns (Fig. 3). Although previous work and the data presented here show that cortical and subcortical sources cannot, in general, be unambiguously resolved, we found that if the distribution of cortical and subcortical sources is sparse, the problem becomes tractable.

These observations motivated us to create a hierarchical SP inverse algorithm to find the set of sparse cortical and subcortical sources that best explain the M/EEG data. Our analyses of various source estimators (Fig. 4) showed that our algorithm has a performance superior to alternatives for the subcortical structures and similar to existing approaches for the cortical structures (36, 42–45).

If the locations of activity in cortical and subcortical structures are known, and each active area can be modeled with an equivalent current dipole, linear least squares can be used to estimate source current time courses (46, 47). However, fitting the locations of multiple dipolar sources usually requires tailored and often interactively guided fitting strategies. Our approach, on the other hand, automatically finds the constellation of sources in a variety of conditions, including those where the source activities may overlap in time. Further, instead of isolated current dipoles we use concise distributed dipole representations within relevant subcortical anatomical subdivisions. Other methods, such as magnetic field tomography (48) and the linearly constrained minimum-variance beamformer (49, 50), that have been applied to locate deep sources implicitly, use some of the principles we formalize here. Our analyses on simulated and experimental data show that algorithms using both sparsity and hierarchy can resolve simultaneously active cortical and subcortical sources under realistic low SNR conditions (Figs. 5 and 6).

Our algorithm relies on being able to specify, at least approximately, the sparsity level $L$ for the cortical and subcortical sources. This parameter can be set based on domain knowledge of the number of expected sources. In this work, we chose conservative values for $L$ exceeding the true or expected level of sparsity and found that the algorithm performed well. In cases where prior domain knowledge does not reasonably exist, formal model selection methods such as cross-validation could be used to choose $L$. Several recent publications have proposed sparsity-based algorithms for M/EEG cortical source estimation (20, 21, 36, 51–61). These and other methods (62) might also be applied in a hierarchical fashion to estimate both cortical and subcortical sources. Further, techniques using distributed sparse representations and dynamical sparsity constraints could be used to estimate subcortical sources in conditions involving more extended areas of cortex (63, 64).

Our analysis implies that sparsity is not merely an assumption, but rather a crucial requirement for cortical and subcortical sources to be jointly estimated. The key question to answer is the extent to which the cortical fields can be distinguished from those of potential subcortical sources. Since the cortical areas that par-

ticipate will vary according to the task, fulfillment of this requirement must be evaluated on case-by-case basis. Notably, even in complex cognitive tasks such as picture naming (23) and viewing or mimicking of lip forms (24), salient activity can be limited to a small number of regions. In each of these regions the source may be focal (well represented by a dipole) or extended (a cortical patch). The methods we have presented in this paper illustrate one approach for characterizing the extent to which subcortical activity can be recovered using M/EEG.

## Materials and Methods

**MRI and M/EEG Acquisition.** We acquired MRI and M/EEG data from four healthy subjects aged $25-45$ y. All subjects provided written informed consent. All studies were approved by the Human Research Committee at Massachusetts General Hospital. Data from one subject were used for the principal angles analysis and simulations. We recorded auditory evoked fields in three other subjects. In one subject, technical problems made the data unusable. For each subject, we obtained T1-weighted structural MRI, a 306-channel MEG, and a 70-electrode EEG recording (*SI Appendix*, section SI I).

**Source Space Construction.** We used FreeSurfer to reconstruct neocortical and hippocampal surfaces and segment subcortical volumes from the MRI (65–67). We placed dipoles with orientations normal to the triangulated surface mesh for neocortex and hippocampus at the gray–white matter interface, with $\sim$1 mm spacing. We placed triplets of orthogonal dipoles in subcortical volumes covering the thalamus, putamen, caudate, brainstem (midbrain), and amygdala, at $\sim$1 mm voxel spacing. To reduce the dimensionality of the source space, we grouped neighboring cortical dipoles into "patches" (68–70). We grouped neighboring subcortical dipoles into "subdivisions" sized to produce signals with similar amplitudes to the cortical patches (*SI Appendix*, section SI I). For cortical patches with an average area of 175 mm$^2$, this sizing procedure resulted in 209 subcortical subdivisions with volumes ranging from 175- to 1,800- mm$^3$ (Fig. 1*B*). Regions with higher current density have finer divisions (higher resolution) than those with low current density (e.g., $\sim$200-mm$^3$ divisions in striatum vs. $\sim$1,795-mm$^3$ divisions in thalamus).

**Forward Solutions.** We derived a three-compartment boundary-element model from the MRI data and numerically computed forward solutions using the MNE software package (68, 69). To account for the different sensor types, units, and noise levels in the M/EEG measurements, we whitened the gain matrices, using an estimate of the observation noise covariance matrix (69). For MEG simulation studies, we set the noise covariance matrix to be similar to typical resting eyes-open recordings, $\mathbf{Q} = \text{diag}([g^2, g^2, m^2, \ldots, g^2, g^2, m^2])$, where $g = 2.5$ fT/cm and $m = 10$ fT. For M/EEG experimental studies, we estimated noise covariance matrices from the resting eyes-open data. To account for differences in current strength across source divisions, we scaled gain matrices for each division by the regional current strengths (71) (*SI Appendix*, section SI I). We constructed the reduced-dimensionality M/EEG gain matrices $\mathbf{G}_k$ for each division $k \in \{\mathcal{B}: 1, 2, \ldots, K\}$, using a singular value decomposition, retaining components to capture >95% of the total spectral energy (36, 70).

**Analysis of Forward Solutions.** We simulated the field pattern $\mathbf{y}_k$ for a division $k$ by activating the most significant eigenmode of $\mathbf{G}_k$ with a unit current. We assessed the degree to which the field pattern $\mathbf{y}_k$ could be explained by some distribution of currents in a region $\mathcal{R} \subset \mathcal{B}$ ($k \notin R$) by fitting $\mathbf{G}_\mathcal{R} \mathbf{x}_\mathcal{R}$ to $\mathbf{y}_k$, using the regularized $l_2$ minimum-norm criterion. We computed $\hat{\mathbf{x}}_\mathcal{R} = \mathbf{G}'_\mathcal{R} \left( \mathbf{G}_\mathcal{R} \mathbf{G}'_\mathcal{R} + \lambda^2 \mathbf{I} \right)^{-1} \mathbf{y}_k$, setting the regularization parameter $\lambda^2$ to 1/9. We then quantified the goodness of fit between the best-fitting field patterns, $\hat{\mathbf{y}}_\mathcal{R} = \mathbf{G}_\mathcal{R} \hat{\mathbf{x}}_\mathcal{R}$ and the original pattern $\mathbf{y}_k$, as $1 - ||\mathbf{y}_k - \hat{\mathbf{y}}_\mathcal{R}||_2/||\mathbf{y}_k||_2$. For visualization of the fields, we dewhitened $\hat{\mathbf{y}}_\mathcal{R}$ and mapped it to a virtual grid of 304 magnetometers distributed evenly across the Elekta helmet (72).

We used principal angles (27, 28) to quantify the correlation between putative field patterns arising from currents in two nonoverlapping regions: $\mathcal{R}_1 \subset \mathcal{B}$ and $\mathcal{R}_2 \subset \mathcal{B}$. We define a configuration of currents in a collection of patches or subdivisions as a combination of eigenmodes from those patches or subdivisions. Any putative field pattern arising from a configuration of currents in $\mathcal{R}_1$ is defined by some subset $\mathbf{U}_{\mathcal{R}_1}$ of the eigenmodes of the gain matrix $\mathbf{G}_{\mathcal{R}_1}$. If $\mathbf{G}_{\mathcal{R}_1}$ comprises $n1$ eigenmodes, there are $2^{n1} - 1$ subsets of eigenmodes $\{\mathbf{U}_{\mathcal{R}_1}\}$ collectively corresponding to all possible current configurations within $\mathcal{R}_1$. Similarly, for region $\mathcal{R}_2$, the gain matrix $\mathbf{G}_{\mathcal{R}_2}$

comprises $n2$ eigenmodes, and there are $2^{n2} - 1$ subsets of eigenmodes $\{\mathbf{U}_{\mathcal{R}_2}\}$ collectively corresponding to all possible current configurations within $\mathcal{R}_2$. The degree to which a field pattern arising from any current configuration in $\mathcal{R}_1$ can be explained by a field pattern arising from any current configuration in $\mathcal{R}_2$ is specified by the set of principal angles $\{\Theta_{1,2}\}$ between each subset of $\{\mathbf{U}_{\mathcal{R}_1}\}$ and each subset in $\{\mathbf{U}_{\mathcal{R}_2}\}$. We computed the set of principal angles $\{\Theta_{1,2}\}$ across all $(2^{n1} - 1)(2^{n2} - 1)$ combinations of subsets of $\{\mathbf{U}_{\mathcal{R}_1}\}$ and $\{\mathbf{U}_{\mathcal{R}_2}\}$.

The principal angles analysis does not account for the gains of the eigenmodes. To ensure that the principal angles histogram is not biased toward smaller eigenmodes, we confined the histogram to include only those subsets of $\{\mathbf{U}_{\mathcal{R}_1}\}$ and $\{\mathbf{U}_{\mathcal{R}_2}\}$ that contained the largest eigenmodes from each patch or subdivision. To illustrate how the principal angles correspond to field patterns with varying levels of similarity, we selected a representative combination of eigenmodes $\mathbf{U}_{\mathcal{R}_1}$ and $\mathbf{U}_{\mathcal{R}_2}$ having angles $35°$, $45°$, $60°$, and $86°$; generated the field $\mathbf{y}_1 = \mathbf{U}_{\mathcal{R}_1}$; projected (minimum $l_2$-norm) $\mathbf{y}_2 = \mathbf{U}_{\mathcal{R}_2}\mathbf{x}_{\mathcal{R}_2}$ to $\mathbf{y}_1$; and plotted the $\mathbf{y}_2$ most similar to $\mathbf{y}_1$.

**Analysis of Resolution Matrices for Inverse Solutions.** We used the resolution matrix to assess performance of the minimum $l_2$-norm (MNE) and SP inverse solutions. The MNE solution for the distributed source space $\mathcal{B}$, with prewhitened gain matrix $\mathbf{G}_{\mathcal{B}}$, is $\mathbf{W}^{\text{MNE}}(\mathbf{G}_{\mathcal{B}}) = \mathbf{R}\mathbf{G}'_{\mathcal{B}}\left(\mathbf{G}_{\mathcal{B}}\mathbf{R}\mathbf{G}'_{\mathcal{B}} + \lambda^2\mathbf{I}\right)^{-1}$.

We specified the prior source covariance $\mathbf{R}$ so that $\text{Tr}(\mathbf{G}_{\mathcal{B}}\mathbf{R}\mathbf{G}'_{\mathcal{B}}) = \text{Tr}(\mathbf{I}) = N$, set $\lambda^2 = 1/9$ (72), and computed the MNE resolution matrix $\mathbf{K}^{\text{MNE}}(\mathbf{G}_{\mathcal{B}}) = \mathbf{W}^{\text{MNE}}(\mathbf{G}_{\mathcal{B}})\mathbf{G}_{\mathcal{B}}$ (30). We used a similar process to compute the MNE resolution matrix $\mathbf{K}^{\text{MNE}}(\mathbf{G}_{\mathcal{B}_r})$. The SP estimates for the source space $\mathcal{B}_r$, with prewhitened gain matrix $\mathbf{G}_{\mathcal{B}_r}$, were obtained using the procedure in *SI Appendix*, section SI III. We characterized the SP resolution matrix $\mathbf{K}^{\text{SP}}$ empirically. We simulated unit currents in the most significant eigenmode of each brain division $i$ and used the gain matrix $\mathbf{G}_{\mathcal{B}_r(i)}$ to generate the noiseless (prewhitened) MEG fields $\mathbf{y}_i$. We then used SP to estimate the source location $q_i = \text{SP}(\mathbf{y}_i, \mathbf{G}_{\mathcal{B}_r}, L = 1)$ and computed the source current $\beta_{\{q_i, i\}}$, using a least-squares fit of $\mathbf{G}_{\mathcal{B}_r(q_i)}$ to $\mathbf{y}_i$. The estimated $q_i$ and $\beta_{\{q_i, i\}}$ together specify the locations and magnitudes of the nonzero elements of the empirical resolution matrix $\mathbf{K}^{\text{SP}}$. As $L = 1$, only the $q_i$th element of column $i$ of $\mathbf{K}^{\text{SP}}$ is nonzero. In the resolution matrices, we ordered the sources according to their physical proximity to each other. Therefore, when sources are estimated accurately, the resolution matrix has a diagonal appearance. We used the resolution matrices $\mathbf{K}^{\text{MNE}}$ and $\mathbf{K}^{\text{SP}}$ to compute the spatial dispersion $\text{SD}_i = \sqrt{\sum_{k=1}^{N}(d_{ki}\cdot||\mathbf{K}_{ki}||)^2/\sum_{k=1}^{N}||\mathbf{K}_{ki}||^2}$ and the dipole localization error $\text{DLE}_i = d_{ji}$, where $j = \arg\max_k\{||\mathbf{K}_{ki}||\}$ and $d_{ji}$ is the distance between centroids of divisions $j$ and $i$ (31).

**Source Estimation Algorithm.** At each hierarchy level, we performed source estimation using the SP algorithm in ref. 36, modified to use relaxed mutual coherence thresholds. The mutual coherence thresholds were set equal to the mean-max correlation among gain matrices from neighborhoods of brain divisions under consideration. Our relaxed mutual coherence thresholds adapt to changing levels of gain matrix correlation across hierarchy levels, giving correlation thresholds in the $\sim0.75-0.90$ range. Additional details are in *SI Appendix*, section SI III, Table S2, and Fig. S4.

**Somatosensory Evoked Response Simulations.** We simulated activity in five regions of interest: a 1-cm$^3$ volume in the left somatosensory thalamus including the ventral posterior area and four $\sim$600- to 800-mm$^2$ surface patches in primary and secondary somatosensory cortices and the posterior parietal area. The simulated cortical current time courses were Gabor atoms of the form $Ae^{-(t-t_0)/2\sigma^2}$, where $A$, $t_0$, and $\sigma$ denote the amplitude, delay, and width of the evoked component and were set based on previous studies (15, 36, 58). The simulated thalamic current time course consisted of 10 repetitions of $\cos^2(2*\pi*f*t + \phi)$ with $f = 100\,\text{Hz}$, $\phi = \pi/3$, duration 15 ms, and repetition period 25 ms. We simulated currents over a time range of $0-250$ ms, with 3-kHz sampling rate. We calculated the MEG signals using the MRI-based forward model and added white Gaussian observation noise with mean zero and variance specified to achieve a SNR of 7 dB.

**Auditory Evoked Response Recordings.** We recorded M/EEG simultaneously while presenting binaural broadband click trains (0.1 ms duration, $65-80$ dB/nHL intensity, 110 ms interstimulus interval, 9.09 Hz click rate) as subjects rested with eyes open (38, 73). We preprocessed the data, performed stimulus-locked averaging to compute the ABR and the MLR, estimated observation noise covariances using the filtered baseline recordings (38) (*SI Appendix*, Fig. S8), whitened the filtered MEG and EEG data using their respective noise covariances, and used both MEG and EEG for source estimation. Further details are in *SI Appendix*, section SI IV.

**Materials Availability.** All code and data from this paper are shared online on our MNE website: martinos.org/mne.

1. Jones EG (2001) The thalamic matrix and thalamocortical synchrony. *Trends Neurosci* 24:595–601.
2. Jones EG (2002) Thalamic circuitry and thalamocortical synchrony. *Philos Trans R Soc Lond B Biol Sci* 357:1659–1673.
3. Steriade M, McCormick DA, Sejnowski TJ (1993) Thalamocortical oscillations in the sleeping and aroused brain. *Science* 262:679–685.
4. Buzsaki G (2016) Theta oscillations in the hippocampus. *Neuron* 33: 325–340.
5. Ching S, Cimenser A, Purdon PL, Brown EN, Kopell NJ (2010) Thalamocortical model for a propofol-induced alpha-rhythm associated with loss of consciousness. *Proc Natl Acad Sci USA* 107:22665–22670.
6. Steriade M, Contreras D, Amzica F, Timofeev I (1996) Synchronization of fast (30-40 hz) spontaneous oscillations in intrathalamic and thalamocortical networks. *J Neurosci* 16:2788–2808.
7. Crunelli V, Hughes SW (2010) The slow rhythm of non-REM sleep: A dialogue between three cardinal oscillators. *Nat Neurosci* 13:9–17.
8. Hughes SW, Crunelli V (2005) Thalamic mechanisms of EEG alpha rhythms and their pathological implications. *Neuroscientist* 11:357–372.
9. Graybiel AM (2000) The basal ganglia. *Curr Biol* 10:509–511.
10. Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* 9:357–381.
11. Haber SN (2003) The primate basal ganglia: Parallel and integrative networks. *J Chem Neuroanat* 26:317–330.
12. Phelps EA, LeDoux JE (2005) Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron* 48:175–187.
13. Douglas RJ, Pribram KH (1966) Learning and limbic lesions. *Neuropsychologia* 4:197–220.
14. Blumenfeld H (2010) *Neuroanatomy Through Clinical Cases* (Sinauer, Sunderland, MA), 2nd Ed.
15. Schomer DL, da Silva FL, eds (2010) *Electroencephalography: Basic Principles, Clinical Applications and Related Fields* (Lippincott Williams & Wilkins, Philadelphia), 6th Ed, pp 669–724.

16. Hämäläinen MS (1993) Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev Mod Phys* 65:413–505.
17. Baillet S, Mosher JC, Leahy RM (2001) Electromagnetic brain mapping. *IEEE Signal Process Mag* 18:14–30.
18. Dale A, Sereno M (1993) Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *J Cogn Neurosci* 5:162–176.
19. Dale AM, et al. (2000) Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* 26:55–67.
20. Ou W, Hämäläinen MS, Golland P (2009) A distributed spatio-temporal EEG/MEG inverse solver. *Neuroimage* 44:932–946.
21. Gramfort A, Strohmeier D, Haueisen J, Hämäläinen MS, Kowalski M (2013) Time-frequency mixed-norm estimates: Sparse M/EEG imaging with non-stationary source activations. *Neuroimage* 70:410–422.
22. Attal Y, Schwartz D (2013) Assessment of subcortical source localization using deep brain activity imaging model with minimum norm operators: An MEG study. *PLoS One* 8:e59856.
23. Salmelin R, Hari R, Lounasmaa OV, Sams M (1994) Dynamics of brain activation during picture naming. *Nature* 368:463–465.
24. Nishitani N, Hari R (2002) Viewing lip forms: Cortical dynamics. *Neuron* 36:1211–1220.
25. Hari R, Ilmoniemi R (1986) Cerebral magnetic fields. *Crit Rev Biomed Eng* 14:93–126.
26. Engemann D, Gramfort A (2015) Automated model selection in covariance estimation and spatial whitening of MEG and EEG signals. *Neuroimage* 108:328–342.
27. Bjorck A, Golub GH (1973) Numerical methods for computing angles between linear subspaces. *Math Comput* 27:579–594.
28. Knyazev AV, Argentati ME (2002) Principal angles between subspaces in an A-based scalar product: Algorithms and perturbation estimates. *SIAM J Sci Comput* 23:2008–2040.
29. Forss N, et al. (1994) Activation of the human posterior parietal cortex by median nerve stimulation. *Exp Brain Res* 99:309–315.

Krishnaswamy et al.

30. Liu AK, Belliveau JW, Dale AM (1998) Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. *Proc Natl Acad Sci USA* 95:8945–8950.

31. Molins A, Stufflebeam SM, Brown EN, Hämäläinen MS (2008) Quantification of the benefit from integrating MEG and EEG data in minimum L2-norm estimation. *Neuroimage* 42:1069–1077.

32. Friedman JH, Tukey J (1973) A projection pursuit algorithm for exploratory data analysis. *IEEE Trans Comput* 23:881–890.

33. Huber P (1985) Projection pursuit. *Ann Stat* 13:435–475.

34. Needell D, Tropp JA (2009) CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Appl Comput Harmonic Anal* 26:301–321.

35. Dai W, Milenkovic O (2009) Subspace pursuit for compressive sensing signal reconstruction. *IEEE Trans Inform Theor* 55:2230–2249.

36. Babadi B, et al. (2014) A subspace pursuit-based iterative greedy hierarchical solution to the neuromagnetic inverse problem. *NeuroImage* 87:427–443.

37. Pockett S (1999) Anesthesia and the electrophysiology of auditory consciousness. *Conscious Cognit* 61:45–61.

38. Parkkonen L, Fujiki N, Makela JP (2009) Sources of auditory brainstem responses revisited: Contribution by magnetoencephalography. *Hum Brain Mapp* 30:1772–1782.

39. Williamson S, Kaufman L (1981) Biomagnetism. *J Magnetism Magn Mater* 22:129–201.

40. Attal Y, et al. (2007) Modeling and detecting deep brain activity with MEG & EEG. *Proc Annu Int Conf IEEE Eng Med Biol Soc* 2007:4937–4940.

41. Hämäläinen M, Hari R (2002) *Brain Mapping, The Methods*, eds Toga AW, Mazziotta JC (Academic, San Diego).

42. Nunez PL (1995) *Neocortical Dynamics and Human EEG Rhythms* (Oxford Univ Press, New York), 1st Ed.

43. Mosher JC, Spencer ME, Leahy RM, Lewis PS (1993) Error bounds for EEG and MEG dipole source localization. *Electroencephalogr Clin Neurophysiol* 86:303–321.

44. Sekihara K, Sahani M, Nagarajan SS (2005) Localization bias and spatial resolution of adaptive and non-adaptive spatial filters for MEG source reconstruction. *Neuroimage* 25:1056–1067.

45. Nunez PL, Srinivasan R (2006) *Electric Fields of the Brain: The Neurophysics of EEG* (Oxford Univ Press, New York), 2nd Ed.

46. Tesche CD (1996) Non-invasive imaging of neuronal population dynamics in human thalamus. *Brain Res* 729:253–258.

47. Tesche CD (1997) Non-invasive detection of ongoing neuronal population activity in normal human hippocampus. *Brain Res* 749:53–60.

48. Ioannides AA, et al. (1995) Magnetic field tomography of cortical and deep processes: Examples of 'real-time mapping' of averaged and single trial MEG signals. *Int J Psychophysiol* 20:161–175.

49. Kensuke Sekihara SSN (2008) *Adaptive Spatial Filters for Electromagnetic Brain Imaging* (Springer, Berlin).

50. Crespo-Garcia M, et al. (2016) Slow-theta power decreases during item-place encoding predict spatial accuracy of subsequent context recall. *Neuroimage* 142:533–543.

51. Gorodnitsky I, George J, Rao B (1995) Neuromagnetic source imaging with FOCUSS: A recursive weighted minimum norm algorithm. *Electroencephalogr Clin Neurophysiol* 95:231–251.

52. Mosher JC, Leahy RM (1998) Recursive music: A framework for EEG and MEG source localization. *IEEE Trans Biomed Eng* 45:1342–1354.

53. Uutela K, Hämäläinen M, Somersalo E (1999) Visualization of magnetoencephalographic data using minimum current estimates. *Neuroimage* 10:173–180.

54. Durka PJ, Matysiak A, Montes EM, Sosa PV, Blinowska KJ (2005) Multichannel matching pursuit and EEG inverse solutions. *J Neurosci Methods* 148:49–59.

55. Friston K, et al. (2008) Multiple sparse priors for the M/EEG inverse problem. *Neuroimage* 39:1104–1120.

56. Vega-Hernandez M, Martinez-Montes E, Sanchez-Bornot J, Lage-Castellanos A, Valdes-Sosa P (2008) Penalized least squares methods for solving the EEG inverse problem. *Stat Sin* 18:1535–1551.

57. Wipf DP, Owen JP, Attias HT, Sekihara K, Nagarajan SS (2010) Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG. *Neuroimage* 49:641–55.

58. Gramfort A, Strohmeier D, Haueisen J, Hamalainen MS, Kowalski M (2011) Functional brain imaging with M/EEG using structured sparsity in time-frequency dictionaries. *Proc Conf Inf Process Med Imaging* 22:600–611.

59. Gramfort A, Kowalski M, Hämäläinen MS (2012) Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods. *Phys Med Biol* 57:1937–1961.

60. Valdes-Sosa P, Vega-Hernandez M, Sanchez-Bornot JM, Martinez-Montes E, Bobes MA (2009) EEG source imaging with spatio-temporal tomographic nonnegative independent component analysis. *Hum Brain Mapp* 30:1898–1910.

61. Vega-Hernandez M, Martinez-Montes E, Sanchez-Bornot J, Lage-Castellanos A, Valdes-Sosa PA (2008) Penalized least squares methods for solving the EEG inverse problem. *Stat Sin* 18:1535–1551.

62. Trujillo-Barreto NJ, Aubert-Vazquez E, Valdes-Sosa PA (2004) Bayesian model averaging in EEG/MEG imaging. *NeuroImage* 21:1300–1319.

63. Babadi B, Kalouptsidis N, Tarokh V (2010) SPARLS: The sparse RLS algorithm. *IEEE Trans Signal Process* 58:4013–4025.

64. Mileounis G, Babadi B, Kalouptsidis N, Tarokh V (2010) An adaptive greedy algorithm with application to nonlinear communications. *IEEE Trans Signal Process* 58:2998–3007.

65. Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage* 9:179–194.

66. Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. *NeuroImage* 9:195–207.

67. Fischl B, et al. (2002) Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron* 33:341–355.

68. Hämäläinen MS, Sarvas J (1989) Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE Trans Biomed Eng* 36:165–171.

69. Gramfort A, et al. (2014) MNE software for processing MEG and EEG data. *NeuroImage* 86:446–460.

70. Limpiti T, Van Veen BD, Wakai RT (2006) Cortical patch basis model for spatially extended neural activity. *IEEE Trans Biomed Eng* 53:1740–1754.

71. Attal Y, Maess B, Friederici A, David O, De C (2012) Head models and dynamic causal modeling of subcortical activity using magnetoencephalographic/electroencephalographic data. *Rev Neurosci* 23:85–95.

72. Hämäläinen MS, Ilmoniemi RJ (1994) Interpreting magnetic fields of the brain: Minimum norm estimates. *Med Biol Eng Comput* 32:35–42.

73. Makela JP, Hamalainen MS, Hari R, McEvoy L (1994) Whole-head mapping of middle-latency auditory evoked magnetic fields. *Electroencephalogr Clin Neurophysiol* 92:414–421.

www.manaraa.com