The AAAI 2017 Spring Symposium on
Science of Intelligence: Computational Principles of Natural and Artificial Intelligence
Technical Report SS-17-07

Sparse Representation Learning Approach Resolves Deep Sources Underlying MEG and EEG Data

Pavitra Krishnaswamy,[1,2,3,4,+] Gabriel Obregon-Henao,[2,6] Sheraz Khan,[2,5] Behtash Babadi,[3] Eugenio Iglesias,[2] Matti S. Hamalainen,[2,5,*] Patrick L. Purdon,[2,6,*]

[2]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston, USA
[3]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, USA
[4]Harvard-MIT Division of Health Sciences and Technology, Cambridge, USA
[5]Department of Radiology and [6]Department of Anesthesia, Harvard Medical School, Boston, USA

Abstract

Several aspects of human intelligence are crucially shaped by neural computations involving subcortical structures located deep in the human brain. As such, non-invasive techniques to characterize subcortical neural currents in the intact, functioning human brain would help elucidate important features of human intelligence. Present techniques for characterizing deep brain activity have limited temporal resolution, and thus cannot resolve millisecond-scale neural currents. Magnetoencephalography (MEG) and electroencephalography (EEG) measure fields generated by neural currents with high temporal resolution. Here, we demonstrate that deep sources underlying M/EEG fields can be resolved with a sparse representation learning approach. We then introduce a novel source estimation algorithm to characterize neural currents within subcortical structures, and demonstrate performance on realistic simulated examples.

Introduction

Several aspects of human intelligence are crucially influenced by deep subcortical structures in the human brain. For example, sensation and perception are shaped by interactions between the thalamus and neocortex; memory is critically influenced by oscillations involving the hippocampus; and learning ability and emotional intelligence are shaped by limbic structures such as the amygdala (Kandel, Schwartz, and Jessell 2000; Nolte 2009). Thus, techniques to non-invasively characterize subcortical neural currents in the intact functioning human brain would help elucidate important features of human intelligence.

However, present techniques have limited spatiotemporal resolution for neural currents deep in the brain. Functional magnetic resonance imaging (fMRI) provides high spatial resolution, but lacks temporal resolution for oscillations or signaling on the millisecond-scale (Logothetis 2008). Magnetoencephalography (MEG) and electroencephalography (EEG) provide high temporal resolution measures of fields arising from neural currents, but do not directly pinpoint specific sources of these fields (Hamalainen 1993; Nunez and Srinivasan 2006). M/EEG source estimation methods are commonly employed to resolve locations of superficial neocortical sources (Mosher et al. 1993; Baillet, Mosher, and Leahy 2001), but it is not clear if they can reliably resolve deeper subcortical sources that generate weak M/EEG signals.

Here, we analyze the geometry of M/EEG fields generated by cortical and subcortical sources to assess the extent to which they can be distinguished from each other. We find that fields from cortical and subcortical sources can be robustly distinguished when spatially sparse representations of the fields are employed. We then introduce an algorithm that employs spatial sparsity priors in a hierarchical procedure to estimate cortical and subcortical source currents underlying M/EEG recordings. Finally, we evaluate performance on evoked response simulations and demonstrate that the algorithm can accurately resolve thalamic and cortical currents in realistic scenarios.

Methods

Problem Setup: Neural current sources, modeled with an ensemble of current dipoles, generate M/EEG fields $\mathbf{Y}$:

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

where $\mathbf{G}$ is the gain matrix determined by Maxwell's equations, $\mathbf{X}$ are the amplitudes of the current dipoles, $\mathbf{V}$ is the noise, N is the number of M/EEG sensors, M is the number of current dipole sources, and T is the number of time points measured. The source estimation problem is to reconstruct $\mathbf{X}$ given $\mathbf{Y}$, $\mathbf{G}$ and the statistics of $\mathbf{V}$ (Hamalainen 1993).

Construction of Source Spaces and Gain Matrices: We acquired high resolution T1-weighted structural MRIs on healthy volunteers (Siemens 3T TimTrio^{TM} scanner, multi-echo MPRAGE sequence) and used them to reconstruct cortical surfaces and segment subcortical volumes (Dale,

Correspondence: +pavitrak@i2r.a-star.edu.sg,
*msh@nmr.mgh.harvard.edu, *patrickp@nmr.mgh.harvard.edu,
* denotes equal contribution

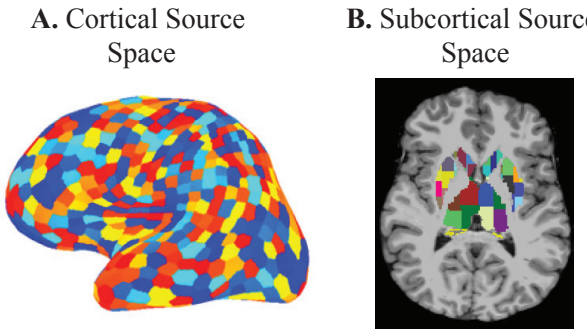

Figure 1: (*A*) Inflated view of cortical source space $\mathcal{C}$: surface patches on the gray-white matter interface. (*B*) 2D slice view of subcortical source space $\mathcal{S}$: subdivided volumes in caudate, putamen, amygdala, thalamus, brainstem, and surface patches on hippocampus.

Fischl, and Sereno 1999; Fischl, Sereno, and Dale 1999; Fischl et al. 2002). We placed dipole sources on cortical surface patches and in subcortical volume subdivisions to form a distributed brain source space $\mathcal{B} = \mathcal{C} \cup \mathcal{S}$ (Figs.1*A*-*B*). For each patch and subdivision in this source space, we numerically computed the gain matrices using the MNE software package (Gramfort et al. 2014). A subset of sources $\mathcal{F} \subset \mathcal{B}$ has gain matrices and source currents denoted by $\mathbf{G}_{\mathcal{F}}$ and $\mathbf{X}_{\mathcal{F}}$ respectively.

Field Geometry Analysis: We consider M/EEG fields $\mathbf{Y}_1$ and $\mathbf{Y}_2$ arising from subsets of sources $\mathcal{D}_1,\ \mathcal{D}_2 \subset \mathcal{B}$ respectively. The degree of distinction between $\mathbf{Y}_1$ and $\mathbf{Y}_2$ is specified by the extent of correlation between subspaces spanned by $\mathbf{G}_{\mathcal{D}_1}$ and $\mathbf{G}_{\mathcal{D}_2}$. Subspace correlations can be quantified using principal angles (Bjorck and Golub 1973). Low angles $\rightarrow 0^\circ$ correspond to overlapping subspaces and indistinguishable fields, while high angles $\rightarrow 90^\circ$ correspond to near-orthogonal subspaces and distinct fields.

Theory and Algorithm

M/EEG Fields due to Cortical vs. Subcortical Currents: We characterized the extent to which M/EEG fields arising from subcortical and cortical sources can be distinguished from each other. We computed principal angles $\Theta_{\mathcal{S},\mathcal{C}}$ between subspaces spanned by the subcortical and cortical gain matrices $\mathbf{G}_{k\,\epsilon\,\mathcal{S}}$ and $\mathbf{G}_{\mathcal{C}}$ (Fig.2, blue). The principal angles across subcortical sources are low, suggesting that fields generated by subcortical sources can be explained by surrogate currents on the cortex.

However, in many neuroscience studies, the set of cortical areas with salient activity at any moment in time is typically not all-encompassing, and instead tends to be limited to a finite number of well-circumscribed areas. Therefore, we considered small subsets of the cortical source space, denoted $\mathcal{C}_{sp}$ and computed principal angles $\Theta_{\mathcal{S},\mathcal{C}_{sp}}$ between subspaces spanned by $\mathbf{G}_{k\,\epsilon\,\mathcal{S}}$ and $\mathbf{G}_{\mathcal{C}_{sp}}$ (Fig.2, brown). The principal angles are now higher, suggesting that fields generated by subcortical sources can be distinguished from those generated by small subsets of cortical sources. We performed

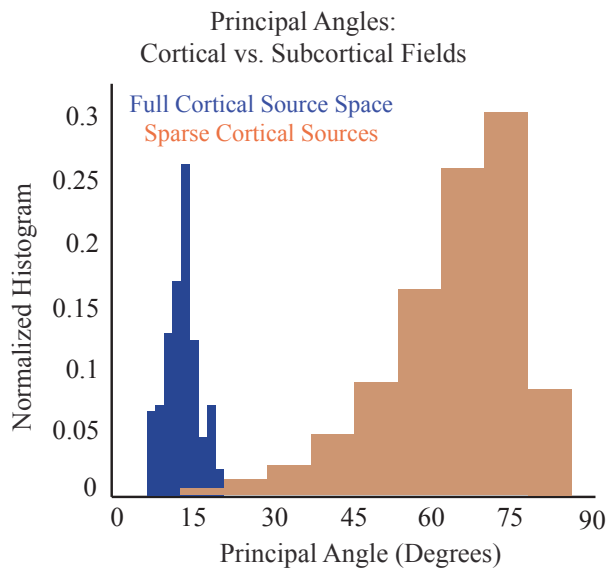

Figure 2: Distribution of principal angles quantifying distinctions between M/EEG fields arising from cortical and subcortical sources. In the presence of the full cortical source space, fields generated by subcortical sources can be sufficiently explained by fields arising from the cortex. However, sparsity in the cortical source space makes it possible to distinguish fields from subcortical and cortical sources.

these analyses for several random subsets of cortical sources and found similar trends, suggesting that employing sparse approximations of $\mathbf{G}_{\mathcal{C}}$ can enable robust distinctions between MEG fields from cortical and subcortical sources.

Source Estimation Algorithm: The above observations suggest that a hierarchical procedure to learn sparse representations of the M/EEG data can enable subcortical source estimation. Suitable algorithms would be (a) effective in identifying relevant sparse features in high-dimensional problems, and (b) amenable to hierarchical 'dictionary learning' type implementations. The class of projection pursuit methods fulfils both criteria.

We employ a subspace pursuit algorithm (Dai and Milenkovic 2009; Needell and Tropp 2009; Babadi et al. 2014) to hierarchically estimate neural currents in both cortical and subcortical structures. Specifically, we first employ subspace pursuit to locate the L sparse cortical sources most relevant to the data, and prune out irrelevant cortical sources. We then construct a composite space of sparse cortical sources and distributed subcortical sources and repeat subspace pursuit to further reduce the space of possible sources. This process systematically refines the distributed brain source space to relevant spatially sparse subsets to reduce dimensionality and improve sensitivity for weak sources, and ultimately outputs locations and time courses of cortical and subcortical source currents underlying the data.

Performance Evaluation

Finally, we demonstrate performance of the algorithm on realistic evoked response simulations. We simulated noisy MEG evoked responses corresponding to a somatosensory median nerve stimulation paradigm (Fig.3*A*). In this paradigm, the ventral posterior lateral thalamus (VPL Th.) is activated during the first $15\,\mathrm{msec}$, followed by activity in the somatosensory cortex. The early MEG fields arising from the VPL Th. are well below the noise floor. We then employed our hierarchical subspace pursuit algorithm to estimate source locations and currents. Fig. 3*B* shows the simulated (dashed) and the estimated (solid) time courses. Fig. 3*C* shows the simulated and estimated source locations. The estimates resemble the simulated ground truth, suggesting the ability to resolve sources in specific thalamic subdivisions.

Conclusions

Our results establish the feasibility of characterizing millisecond-scale neural currents within subcortical regions using non-invasive M/EEG recordings in humans. As such, this work opens up unique opportunities to study aspects of human intelligence – such as perception, emotion and memory – that involve subcortical structures and have a significant influence on higher cortical reasoning.

Acknowledgments

We acknowledge helpful discussions on canonical correlations and sparsity with Emery N. Brown at the Massachusetts Institute of Technology. This work was supported by NIH grants P41-EB015896, 5R01-EB009048, 1S1-0RR031599-01 and NCRR grant 1S10RR031599-01 to M.S.H, and NIH grant DP2-OD006454 to P.L.P.

References

Babadi, B.; Obregon-Henao, G.; Lamus, C.; Hämäläinen, M. S.; Brown, E. N.; and Purdon, P. L. 2014. A subspace pursuit-based iterative greedy hierarchical solution to the neuromagnetic inverse problem. *NeuroImage* 87:427–43.

Baillet, S.; Mosher, J. C.; and Leahy, R. M. 2001. Electromagnetic brain mapping. *IEEE Signal Processing Magazine* 14–30.

Bjorck, A., and Golub, G. H. 1973. Numerical methods for computing angles between linear subspaces. *Mathematics of Computation* 27(123):579–594.

Dai, W., and Milenkovic, O. 2009. Subspace pursuit for compressive sensing signal reconstruction. *IEEE Transactions on Information Theory* 55(5):2230–2249.

Dale, A. M.; Fischl, B.; and Sereno, M. I. 1999. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage* 9:179–194.

Fischl, B.; Salat, D. H.; Busa, E.; Albert, M.; Dieterich, M.; Haselgrove, C.; van der Kouwe, A.; Killiany, R.; Kennedy, D.; Klaveness, S.; Montillo, A.; Makris, N.; Rosen, B.; and Dale, A. M. 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33(3):341–55.

Fischl, B.; Sereno, M. I.; and Dale, A. M. 1999. Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. *NeuroImage* 9:195–207.

Gramfort, A.; Luessi, M.; Larson, E.; Engemann, D. A.; Strohmeier, D.; Brodbeck, C.; Parkkonen, L.; and Hämäläinen, M. S. 2014. MNE software for processing MEG and EEG data. *NeuroImage* 86:446–60.

Hamalainen, M. S. 1993. Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics* 65(2):413–505.

Kandel, E.; Schwartz, J.; and Jessell, T. 2000. *Principles of Neural Science*. McGraw-Hill Companies, fifth edition.

Logothetis, N. K. 2008. What we can do and what we cannot do with fMRI. *Nature* 453(7197):869–78.

Mosher, J. C.; Spencer, M. E.; Leahy, R. M.; and Lewis, P. S. 1993. Error bounds for EEG and MEG dipole source localization. *Electroencephalography and Clinical Neurophysiology* 86(5):303–21.

Needell, D., and Tropp, J. A. 2009. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis* 26(3):301–321.

Nolte, J. 2009. *The Human Brain: An Introduction to its Functional Anatomy*. Mosby/Elsevier, sixth (illustrated) edition.

Nunez, P. L., and Srinivasan, R. 2006. *Electric Fields of the Brain: the Neurophysics of EEG*. Oxford University Press, USA, second edition.

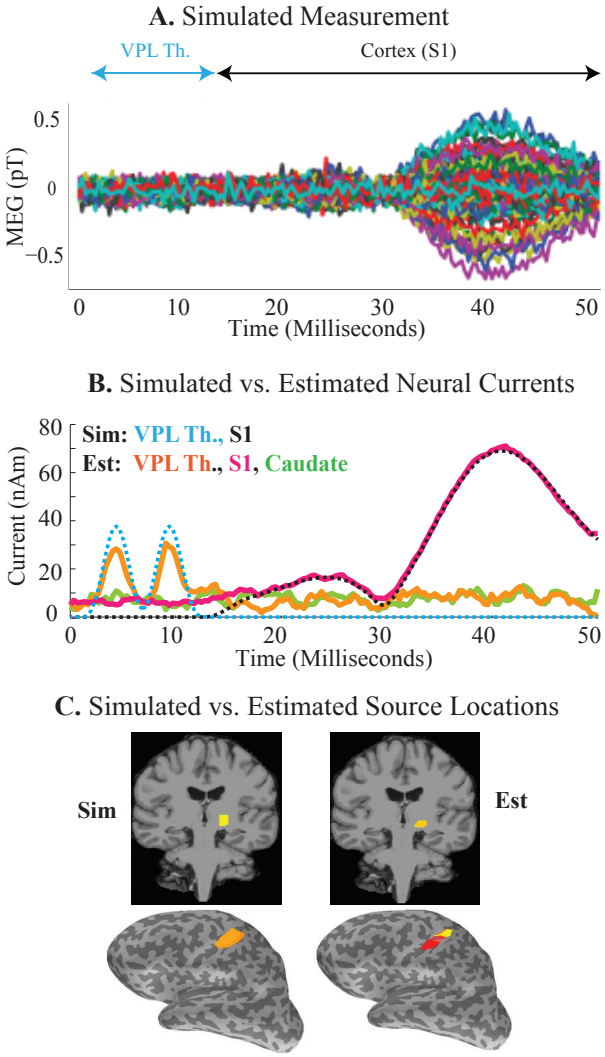

Figure 3: (*A*) Simulated MEG evoked responses corresponding to a somatosensory median nerve stimulation paradigm (one trace per sensor). (*B*) Simulated (dashed) vs. estimated (solid) time courses for neural currents. (*C*) Source locations estimated (right) by hierarchical subspace pursuit closely approximate those in simulations (left).