Algorithm Details: Tables S2 and S3

For hierarchy j

Do subspace pursuit: $[\mathcal{H}_j, \hat{\mathbf{X}}_{\mathcal{H}_j}] = \mathrm{SP}\left(\mathbf{Y}, \mathbf{G}^{(j)}, L\right)$

Initialization:

1. Compute the typical maximum correlation μ between gain matrices of neighboring brain divisions in $\mathbf{G}^{(j)}$ (SUPP).
2. Support: $\mathcal{H}(0)$ $=\{L$ rows in $\hat{\mathbf{X}}^{\mathrm{MNE}}(\mathbf{G}^{(j)}, \mathbf{Y})$ that have the largest l_2-norms across time and satisfy $C(\mathbf{G}^{(j)}_{\mathcal{H}(0)}) < \mu\}$.
3. Residual: $\mathbf{F}(0) = \mathbf{Y} - \mathbf{G}^{(j)}_{\mathcal{H}(0)}\hat{\mathbf{X}}^{\mathrm{MNE}}(\mathbf{G}^{(j)}_{\mathcal{H}(0)}, \mathbf{Y})$.

Iteration from $l = 1$:

1. Additional support: $\mathcal{H}^{\mathrm{ex}}(l) = \{L$ rows in $\hat{\mathbf{X}}^{\mathrm{MNE}}(\mathbf{G}^{(j)}, \mathbf{F}(l-1))$ that have the largest l_2-norms across time and satisfy $C(\mathbf{G}^{(j)}_{\mathcal{H}^{\mathrm{ex}}(l)}) < \mu\}$.
2. Support expansion: $\mathcal{H}(l) = \mathcal{H}(l-1) \cup \mathcal{H}^{\mathrm{ex}}(l)$.
3. Support trimming: $\mathcal{H}(l)$ $=\{L$ rows in $\hat{\mathbf{X}}^{\mathrm{MNE}}(\mathbf{G}^{(j)}_{\mathcal{H}(l)}, \mathbf{Y})$ that have the largest l_2-norms across time and satisfy $C(\mathbf{G}^{(j)}_{\mathcal{H}(l)}) < \mu\}$.
4. Residual update: $\mathbf{F}(l) = \mathbf{Y} - \mathbf{G}^{(j)}_{\mathcal{H}(l)}\hat{\mathbf{X}}^{\mathrm{MNE}}(\mathbf{G}^{(j)}_{\mathcal{H}(l)}, \mathbf{Y})$.

Stop iterations if $\mathcal{H}(l) = \mathcal{H}(l-1)$ and $\mathbf{F}(l) \leq \mathbf{F}(l-1)$.

Outputs: $\mathcal{H}_j = \mathcal{H}(l) \subset \mathcal{B}(j)$ and $\hat{\mathbf{X}}_{\mathcal{H}_j} = \hat{\mathbf{X}}^{\mathrm{MNE}}(\mathbf{G}^{(j)}_{\mathcal{H}(l)}, \mathbf{Y})$.

If $j \neq J$, use $\mathcal{H}_j$to define $\mathcal{B}(j+1)$ and in turn $\mathbf{G}^{(j+1)}$. Else, stop and output joint cortical and subcortical source estimates $\mathcal{H}_J$ and $\hat{\mathbf{X}}_{\mathcal{H}_J}$.

Test Case	Measurements	Cortical Patch Decomposition	Max. Mutual Coherence	Min. Orthogonality Requirement (Degrees)
Figure 4 Resolution	MEG	Hybrid Deep and C-3	0.85	31.8
Somatosensory Evoked Potential	MEG	C-1	0.77	39.6
Somatosensory Evoked Potential	MEG	C-2	0.85	31.8
Somatosensory Evoked Potential	MEG	C-3	0.88	28.4
Somatosensory Evoked Potential	MEG	Hybrid Deep and C-3	0.85	31.8
Auditory Evoked Response	MEG-EEG	C-1 (MLR)	0.85	31.8
Auditory Evoked Response	MEG-EEG	C-2 (MLR)	0.89	27.1
Auditory Evoked Response	MEG-EEG	C-3 (MLR)	0.91	24.5
Auditory Evoked Response	MEG-EEG	Hybrid Deep and C-3 (HFABR)	0.92	23.1

Table 2. Mutual Coherence Thresholds for Subspace Pursuit. The thresholds were computed using whitened modes in each case, such that units were consistent across magnetometers, gradiometers and EEG. When MEG and EEG were considered jointly, the columns of the whitened EEG forward matrix and whitened MEG forward matrices were concatenated (across sensors) for each patch of interest. Forward solutions from neighboring patches are less correlated for coarser patches than for finer patches, thus the maximum coherence allowed increases with increasing fineness. This causes the search to enforce greater orthogonality in coarser source spaces allowing the algorithm to search widely before settling in on relevant regions. Finer source spaces allow more clustering, as the solutions are getting narrowed into the relevant regions.