

Towards human SuperEEG

Lucy L. W. Owen¹, Andrew C. Heusser^{1,2}, and Jeremy R. Manning^{1*}

¹Department of Psychological and Brain Sciences, Dartmouth College,
Hanover, NH 03755, USA
²Akili Interactive,
Boston, MA 02110, USA

Abstract

Human *SuperEEG*¹ entails measuring ongoing neural activity with perfect precision and at arbitrarily high spatiotemporal resolution. Although true SuperEEG is impossible using existing methods, here we present a model-based method for *inferring* neural activity at millimeter-scale spatial resolutions and millisecond-scale temporal resolutions using standard human intracranial recordings. Our approach assumes that different people’s brains exhibit similar spatial correlations, and that (all else being equal) neural activity at nearby locations will tend to be similar. One can then ask, for an arbitrary individual’s brain: given recordings from a limited set of locations in that individual’s brain, along with the observed spatial correlations in other people’s recordings, what would recordings most likely have looked like at *other* locations in that individual’s brain?

Keywords: Electrocorticography (ECoG), intracranial electroencephalography (iEEG), local field potential (LFP), epilepsy, maximum likelihood estimation, Gaussian process regression

Introduction

Modern human brain recording techniques are fraught with compromise [2]. Commonly used approaches include functional magnetic resonance imaging (fMRI), scalp electroencephalography (EEG), and magnetoencephalography (MEG). For each of these techniques, neuroscientists and electrophysiologists must choose to optimize spatial resolution at the cost of temporal resolution (e.g., as in fMRI) or temporal resolution at the cost of spatial resolution (e.g., as in EEG and MEG). A less widely used approach (due to requiring work with neurosurgical patients) is to record from electrodes implanted directly onto the cortical surface (electrocorticography; ECoG) or into deep brain structures (intracranial EEG; iEEG). However, these intracranial approaches also require compromise: the high temporal and spatial resolu-

¹The term “SuperEEG” was coined by Robert J. Sawyer in his popular science fiction novel *The Terminal Experiment* [1]

25 tions of intracranial recordings comes at the cost of substantially reduced brain coverage, since safety
26 considerations limit the number of electrodes one may implant in a given patient’s brain. Further, the
27 locations of implanted electrodes are determined by clinical, rather than research, needs.

28 An increasingly popular approach is to improve the effective spatial resolution of MEG or scalp
29 EEG data by using a geometric approach called *beamforming* to solve the biomagnetic or bioelectrical
30 inverse problem [3]. This approach entails using detailed brain conductance models (often informed
31 by high spatial resolution anatomical MRI images) along with the known sensor placements (localized
32 precisely in 3D space) to reconstruct brain signals originating from theoretical point sources deep in the
33 brain (and far from the sensors). Traditional beamforming approaches must overcome two obstacles.
34 First, the inverse problem beamforming seeks to solve has infinitely many solutions. Researchers have
35 made traction towards constraining the solution space by assuming that signal-generating sources are
36 localized on a regularly spaced grid spanning the brain and that individual sources are small relative to
37 their distances to the sensors [4–6]. The second, and in some ways much more serious, obstacle is that
38 the magnetic fields produced by external (noise) sources are substantially stronger than those produced
39 by the neuronal changes being sought (i.e., at deep structures, as measured by sensors at the scalp). This
40 means that obtaining adequate signal quality often requires averaging the measured responses over tens
41 to hundreds of responses or trials (e.g., see review by [6]).

42 Another approach to obtaining high spatial and temporal resolution neural data has been to collect
43 fMRI and EEG data simultaneously. Simultaneous fMRI-EEG has the potential to balance the high spa-
44 tial resolution of fMRI with the high temporal resolution of scalp EEG, thereby, in theory, providing the
45 best of both worlds. In practice, however, the signal quality of both recordings suffers substantially when
46 the two techniques are applied simultaneously (e.g., see review by [7]). In addition, the experimental
47 designs that are ideally suited to each technique individually are somewhat at odds. For example, fMRI
48 experiments typically lock stimulus presentation events to the regularly spaced image acquisition time
49 (TR), which maximizes the number of post-stimulus samples. By contrast, EEG experiments typically

50 employ jittered stimulus presentation times to maximize the experimentalist’s ability to distinguish elec-
51 trical brain activity from external noise sources such as from 60 Hz alternating current power sources.

52 The current “gold standard” for precisely localizing signals and sampling at high temporal resolution
53 is to take (ECoG or iEEG) recordings from implanted electrodes (but from a limited set of locations in
54 any given brain). This begs the following question: what can we infer about the activity exhibited by
55 the rest of a person’s brain, given what we learn from the limited intracranial recordings we have from
56 their brain and additional recordings taken from *other* people’s brains? Here we develop an approach,
57 which we call *SuperEEG*, based on Gaussian process regression [8]. SuperEEG entails using data from
58 multiple people to estimate activity patterns at arbitrary locations in each person’s brain (i.e., independent
59 of their electrode placements). We test SuperEEG approach using two large datasets of intracranial
60 recordings [9–22]. We show that the SuperEEG algorithm recovers signals well from electrodes that were
61 held out of the training dataset. We also examine the factors that influence how accurately activity may
62 be estimated (recovered), which may have important implications for electrode design and placement in
63 neurosurgical applications.

64 Approach

65 The SuperEEG approach to inferring high temporal resolution full-brain activity patterns is outlined and
66 summarized in Figure 1. We describe (in this section) and evaluate (in *Results*) our approach using a two
67 large previously collected dataset comprising multi-session intracranial recordings. Dataset 1 comprises
68 multi-session recordings taken from 6876 electrodes implanted in the brains of 88 epilepsy patients [9–
69 13]. Each recording session lasted from XXX–XXX hours, and includes data recorded roughly from
70 when the patients woke up each morning, to before they went to sleep at the end of each day. In addition
71 to typical bed-ridden hospital patient activities (e.g., lying in bed, reading, watching television, using
72 personal electronic devices, listening to music, visiting with family and friends, etc.), the patients also
73 performed a variety of experimental cognitive tasks throughout their day (primarily list-learning memory

74 tasks). For the purposes of the Dataset 1 analyses presented here, we aggregated all data across each
75 recording session, ignoring the particular activities or tasks the patients were performing at any given
76 moment. We used Dataset 1 to develop and debug our main SuperEEG approach, and to examine the
77 extent to which SuperEEG might be able to generate task-general predictions. Dataset 2 comprised
78 multi-session recordings from XXX electrodes implanted in the brains of XXX epilepsy patients [14–
79 22]. Whereas Dataset 1 included recordings taken during a wide variety of behaviors, Dataset 2 included
80 recordings taken as each patient performed each of two memory tasks: a random word list free recall
81 task and a categorized word list free recall task. We used Dataset 2 to further examine the ability of
82 SuperEEG to generalize its predictions within versus across tasks.

83 We first applied fourth order Butterworth notch filter to remove 60 Hz ($\pm .5$ Hz) line noise from
84 every recording (from every electrode). Next, we downsampled the recordings (regardless of the original
85 samplerate) to 250 Hz. (This downsampling step served to both normalize for differences in sampling
86 rates across patients and to ease the computational burden of our subsequent analyses.) We then excluded
87 any electrodes that showed putative epileptiform activity. Specifically, we excluded from further analysis
88 any electrode that exhibited an average kurtosis of 10 or greater across all of that patient’s recording
89 sessions. We also excluded any patients with fewer than 2 electrodes that passed this criteria, as the
90 SuperEEG algorithm requires measuring correlations between 2 or more electrodes from each patient.
91 For Dataset 1, this yielded clean recordings from 4168 electrodes implanted throughout the brains of 67
92 patients (Fig. 1A); for Dataset 2, this yielded clean recordings from 2975 electrodes from 24 patients.
93 Each individual patient contributes electrodes from a limited set of brain locations, which we localized in
94 a common space [MNI152; 24]; an example Dataset 1 patient’s 54 electrodes that passed the predefined
95 kurtosis test are highlighted in black and red.

The recording from a given electrode is maximally informative about the activity of the neural tissue
immediately surrounding its recording surface. However, brain regions that are distant from the recording
surface of the electrode also contribute to the recording, albeit (*ceteris paribus*) to a much lesser extent.

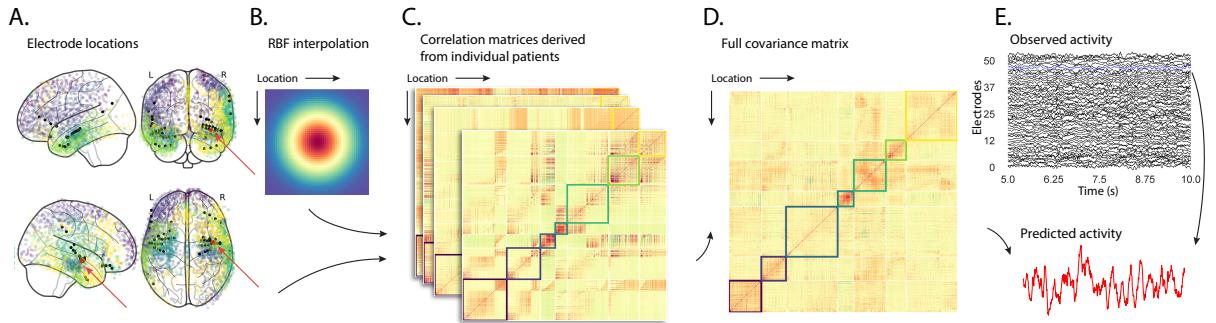


Figure 1: Methods overview. **A. Electrode locations.** Each dot reflects the location of a single electrode implanted in the brain of a Dataset 1 patient. A held-out recording location from one patient is indicated in red, and the patient’s remaining electrodes are indicated in black. The electrodes from the remaining patients are colored by k -means cluster (computed using the full-brain correlation model shown in Panel D). **B. Radial basis function kernel.** Each electrode contributed by the patient (black) weights on the full set of locations under consideration (all dots in Panel A, defined as \bar{R} in the text). The weights fall off with positional distance (in MNI space) according to an RBF. **C. Per-patient correlation matrices.** After computing the pairwise correlations between the recordings from each patient’s electrodes, we use RBF-weighted averages to estimate correlations between all locations in \bar{R} . We obtain an estimated full-brain correlation matrix using each patient’s data. **D. Merged correlation model.** We combine the per-patient correlation matrices (Panel C) to obtain a single full-brain correlation model that captures information contributed by every patient. Here we have sorted the rows and columns to reflect k -means clustering labels [using $k=7$; 23], whereby we grouped locations based on their correlations with the rest of the brain (i.e., rows of the matrix displayed in the panel). The boundaries denote the cluster groups. The rows and columns of Panel C have been sorted using the Panel D-derived cluster labels. **E. Reconstructing activity throughout the brain.** Given the observed recordings from the given patient (shown in black; held-out recording is shown in blue), along with a full-brain correlation model (Panel D), we use Equation 12 to reconstruct the most probable activity at the held-out location (red).

One mechanism underlying these contributions is volume conduction. The precise rate of falloff due to volume conduction (i.e., how much a small volume of brain tissue at location x contributes to the recording from an electrode at location η) depends on the size of the recording surface, the electrode's impedance, and the conductance profile of the volume of brain between x and η . As an approximation of this intuition, we place a Gaussian radial basis function (RBF) at the location η of each electrode's recording surface (Fig. 1B). We use the values of the RBF at any brain location x as a rough estimate of how much structures around x contributed to the recording from location η :

$$\text{rbf}(x|\eta, \lambda) = \exp \left\{ -\frac{\|x - \eta\|^2}{\lambda} \right\}, \quad (1)$$

where the width variable λ is a parameter of the algorithm (which may in principle be set according to location-specific tissue conductance profiles) that governs the level of spatial smoothing. In choosing λ for the analyses presented here, we sought to maximize spatial resolution (which implies a small value of λ) while also maximizing the algorithm's ability to generalize to any location throughout the brain, including those without dense electrode coverage (which implies a large value of λ). Using our prior work as a guide [25, 26], we set $\lambda = 20$, although this could in theory be optimized, e.g., using cross validation or a formal model [e.g., 26].

A second mechanism whereby a given region x can contribute to the recording at η is through anatomical connections between structures near x and η . We use spatial correlations in the data to estimate these anatomical connections [27]. Let \bar{R} be the set of locations at which we wish to estimate local field potentials, and let R_s be set of locations at which we observe local field potentials from patient s (excluding the electrodes that did not pass the kurtosis test described above). In the analyses below we define $\bar{R} = \cup_{s=1}^S R_s$. We can calculate the expected inter-electrode correlation matrix for patient s , where $C_{s,k}(i, j)$ is the correlation between the time series of voltages for electrodes i and j from subject s during session k , using:

$$\bar{C}_s = r\left(\frac{1}{n}\left(\sum_{k=1}^n z(C_{s,k})\right)\right), \text{ where} \quad (2)$$

$$z(r) = \frac{\log(1+r) - \log(1-r)}{2} \text{ is the Fisher } z\text{-transformation and} \quad (3)$$

$$z^{-1}(z) = r(z) = \frac{\exp(2z) - 1}{\exp(2z) + 1} \text{ is its inverse.} \quad (4)$$

111 Next, we use Equation 1 to construct a number of to-be-estimated locations by number of patient elec-
 112 trode locations weight matrix, W_s . Specifically, W_s approximates how informative the recordings at
 113 each location in R_s are in reconstructing activity at each location in \bar{R} , where the contributions fall off
 114 with an RBF according to the distances between the corresponding locations:

$$W_s(i,j) = \text{rbf}(i|j, \lambda). \quad (5)$$

115 Given this weight matrix, W_s , and the observed inter-electrode correlation matrix for patient s , \bar{C}_s ,
 116 we can estimate the correlation matrix for all locations in \bar{R} (\hat{C}_s ; Fig. 1C) using:

$$\hat{N}_s(x, y) = \sum_{i=1}^{|R_s|} \sum_{j=1}^{i-1} W(x, i) \cdot W(y, j) \cdot z(\bar{C}_s(i, j)) \quad (6)$$

$$\hat{D}_s(x, y) = \sum_{i=1}^{|R_s|} \sum_{j=1}^{i-1} W(x, i) \cdot W(y, j). \quad (7)$$

$$\hat{C}_s = r\left(\frac{\hat{N}_s}{\hat{D}_s}\right). \quad (8)$$

After estimating the numerator (\hat{N}_s) and denominator (\hat{D}_s) placeholders for each \hat{C}_s , we aggregate these estimates across patients to obtain a single expected full-brain correlation matrix (\hat{K} ; Fig. 1D):

$$\hat{K} = r\left(\frac{\sum \hat{N}_s}{\sum \hat{D}_s}\right). \quad (9)$$

117 Intuitively, the numerators capture the general structures of the patient-specific estimates of full-brain
 118 correlations, and the denominators account for which locations were near the implanted electrodes in

119 each patient. To obtain \hat{K} , we compute a weighted average across the estimated patient-specific full-
 120 brain correlation matrices, where patients with observed electrodes near a particular set of locations in
 121 \hat{K} contribute more to the estimate.

122 Having used the multi-patient data to estimate a full-brain correlation matrix at the set of locations
 123 in \bar{R} that we wish to know about, we next use \hat{K} to estimate activity patterns everywhere in \bar{R} , given
 124 observations at only a subset of locations in \bar{R} (Fig. 1E).

125 Let α_s be the set of indices of patient s 's electrode locations in \bar{R} , and let β_s be the set of indices
 126 of all other locations in \bar{R} . In other words, β_s reflects the locations in \bar{R} where we did not observe a
 127 recording for patient s (these are the recording locations we will want to fill in using SuperEEG). We can
 128 sub-divide \hat{K} as follows:

$$\hat{K}_{\beta_s, \alpha_s} = \hat{K}(\beta_s, \alpha_s), \text{ and} \quad (10)$$

$$\hat{K}_{\alpha_s, \alpha_s} = \hat{K}(\alpha_s, \alpha_s). \quad (11)$$

129 Here $\hat{K}_{\beta_s, \alpha_s}$ represents the correlations between the “unknown” activity at the locations in β_s and the
 130 observed activity at the locations in α_s , and $\hat{K}_{\alpha_s, \alpha_s}$ represents the correlations between the observed
 131 recordings (at the locations in α_s).

132 Let Y_{s,k,α_s} be the number-of-timepoints (T) by $\text{length}(\alpha_s)$ matrix of (observed) voltages from the
 133 electrodes in α_s during session k from patient s . Then we can estimate the voltage from patient s 's k^{th}
 134 session at the locations in β_s using [8]:

$$\hat{Y}_{s,k,\beta_s} = ((\hat{K}_{\beta_s, \alpha_s} \cdot \hat{K}_{\alpha_s, \alpha_s}^{-1}) \cdot Y_{s,k,\alpha_s}^T)^T. \quad (12)$$

135 This equation is the foundation of the SuperEEG algorithm. Whereas we observe recordings only at the
 136 locations in α_s , Equation 12 allows us to estimate the recordings at all locations in β_s , which we can

137 define *a priori* to include any locations we wish, throughout the brain. This yields estimates of the time-
138 varying voltages at *every* location in \bar{R} , provided that we define \bar{R} in advance to include the union of all
139 of the locations in α_s and all of the locations at which we wish to estimate recordings (i.e., a timeseries
140 of voltages).

141 We designed our approach to be agnostic to electrode impedances, as electrodes that do not exist
142 do not have impedances. Therefore our algorithm recovers voltages in standard deviation (z -scored)
143 units rather than attempting to recover absolute voltages. (This property reflects the fact that $\hat{K}_{\beta_s, \alpha_s}$ and
144 $\hat{K}_{\alpha_s, \alpha_s}$ are correlation matrices rather than covariance matrices.) Also, note that Equation 12 directly
145 requires computing a T by T matrix, which can become computationally intractable when T is very
146 large (e.g., for the patient highlighted in Fig. 2, $T = 20458799$). However, because Equation 12 is time
147 invariant, we may compute Y_{s,k,β_s} in a piecewise manner by filling in Y_{s,k,β_s} one row at a time (using
148 the corresponding samples from Y_{s,k,α_s}).

149 The SuperEEG algorithm described above and in Figure 1 allows us to estimate, up to a constant
150 scaling factor, local field potentials (LFPs) for each patient at all arbitrarily chosen locations in the set
151 \bar{R} , even if we did not record that patient's brain at all of those locations. We next turn to an evaluation
152 of the accuracy of those estimates.

153 Results

154 We used a cross-validation approach to test the accuracy with which the SuperEEG algorithm recon-
155 structs activity throughout the brain. For each patient in turn, we estimated full-brain correlation matrices
156 (Eqn. 9) using data from all of the *other* patients. This step ensured that the data we were reconstruct-
157 ing could not also be used to estimate the between-location correlations that drove the reconstructions
158 via Equation 12 (otherwise the analysis would be circular). For that held-out patient, for each of their
159 electrodes in turn, we used Equation 12 to reconstruct activity at the held-out electrode location, using
160 the correlation matrix trained on all other patients' data at \hat{K} , and using activity recorded from the other

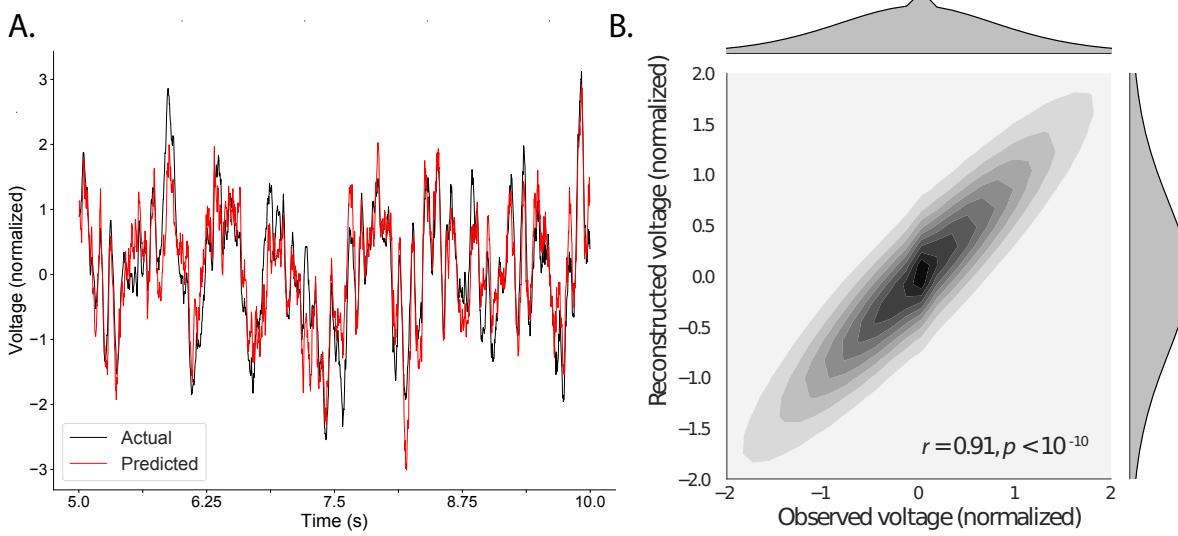


Figure 2: Observed and reconstructed LFP from a single electrode. A. Example LFP. A 5 s recording from the red electrode in Figure 1A is displayed in red, and the reconstructed LFP during the same time window is shown in blue. **B. Observed versus reconstructed LFP over 14.2 hours.** The two-dimensional histogram reflects the relation between distributions of observed versus reconstructed voltages from one patient, across the 14.2 hours of recorded data collected over 6 recording sessions. The correlation reported in the panel is between the observed and reconstructed voltages. Both panels: all voltages are represented in standard deviation units (computed within session).

161 electrodes from the held-out patient as Y_{s,k,α_s} . We then asked: how closely did each of the SuperEEG-
 162 estimated recordings at those electrodes match the observed recordings from those electrodes (i.e., how
 163 closely did the estimated \hat{Y}_{s,k,β_s} match the observed Y_{s,k,β_s} ?).

164 To illustrate our approach, we first examine and individual held-out raw LFP trace and its associ-
 165 ated SuperEEG-derived reconstruction. Figure 2A displays the observed LFP from the red electrode in
 166 Figure 1A (red), and its associated reconstruction (blue), during a 5 s time window during one of the
 167 example patient’s six recording sessions. The two traces match closely ($r = XXX, p = XXX$). Fig-
 168 ure 2B displays a two-dimensional histogram of the actual versus reconstructed voltages for the entire
 169 14.2 total hours of recordings from the example electrode (correlation: $r = XXX, p = XXX$). This
 170 example confirms that the SuperEEG algorithm recovers the recordings from this single electrode well.
 171 Next, we quantify the algorithm’s performance across the full dataset.

172 For each held-out electrode, from each held-out patient in turn, we computed the average corre-
173 lation (across recording sessions) between the SuperEEG-reconstructed voltage traces and the observed
174 voltage traces from that electrode. For this analysis we set \bar{R} to be the union of all electrode locations
175 across all patients. This yielded a single correlation coefficient for each electrode location in \bar{R} , reflecting
176 how well the SuperEEG algorithm was able to recover the recording at that location by incorporating data
177 across patients (black histogram in Fig. 3A, map in Fig. 3C). The observed distribution of correlations
178 was centered well above zero (mean: XXX; t -test comparing mean of distribution of z -transformed per-
179 electrode correlation coefficients to 0: $t(XXX) = XXX, p = XXX$), indicating that the SuperEEG
180 algorithm recovers held-out activity patterns substantially better than random guessing.

181 As a stricter benchmark, we compared the quality of these across-participant reconstructions (i.e.,
182 computed using a correlation model derived from other patients' data) to reconstructions generated using
183 a correlation model trained using the in-patient's data. In other words, for this within-patient benchmark
184 analysis we estimated \hat{C}_s (Eqn. 8) for each patient in turn, using recordings from all of that patient's
185 electrodes except at the location we were reconstructing. These within-patient reconstructions serve as
186 an estimate of how well data from all of the other electrodes from a single patient may be used to esti-
187 mate held-out data. This allows us to ask how much information about the activity at a given electrode
188 might be inferred through (a) volume conductance or other sources of "leakage" from activity patterns
189 measured from the patient's other electrodes and (b) across-electrode correlations learned from that sin-
190 gle patient. As shown in Figure 3A (gray histogram), the distribution of within-patient correlations was
191 centered well above zero (mean: XXX; t -test comparing mean of distribution of z -transformed per-
192 electrode correlation coefficients to 0: $t(XXX) = XXX, p = XXX$). However, the across-patient
193 correlations were substantially higher (t -test comparing average z -transformed within versus across pa-
194 tient electrode correlations: $t(XXX) = XXX, p = XXX$). This is an especially conservative test,
195 given that the across-patient SuperEEG reconstructions exclude (from the correlation matrix estimates)
196 all data from the patient whose data is being reconstructed. We repeated each of these analyses on a

197 second independent dataset and found similar results (Fig. 3B, D; within versus across reconstruction
198 accuracy: $t(23) = 6.93, p < 10^{-5}$). We also replicated this result separately for each of the two ex-
199 periments from Dataset 2 (Fig. S2). This overall finding, that when reconstructing held-out data from a
200 given patient correlation models derived from *other* patient's data yield higher reconstruction accuracy
201 than correlation models derived from that patient, has two important implications. First, it implies that
202 distant electrodes provide additional predictive power to the data reconstructions beyond the information
203 contained solely in nearby electrodes. (This follows from the fact that each patient's electrodes are im-
204 planted in a unique set of locations, so for any given electrode the closest electrodes in the full dataset
205 are likely to come from the same patient.) Second, it implies that the spatial correlations derived from
206 the SuperEEG algorithm are, to some extent, similar across people.

207 The recordings we analyzed from Dataset 1 comprised data collected as the patients performed a
208 variety of (essentially uncontrolled) tasks throughout each day's recording session. That we observed
209 reliable reconstruction accuracy across patients suggests that the spatial correlations derived from the
210 SuperEEG algorithm are, to some extent, similar across tasks. We tested this finding more explicitly
211 using Dataset 2. In Dataset 2, the recordings were limited to times when each patient was participating
212 in each of two experiments (Experiment 1, a random-word list free recall task, and Experiment 2, a
213 categorized list free recall task). We wondered whether a correlation model trained using data only from
214 one experiment might yield good predictions for data from the other experiment. Further, we wondered
215 about the extent to which it might be beneficial or harmful to combine data across tasks.

216 To test the task-specificity of the SuperEEG-derived correlation models, we repeated our within-
217 and across-patient cross validation procedures separately for Experiment 1 and Experiment 2 data from
218 Dataset 2. We then compared the reconstruction accuracies of held-out electrodes, for models trained
219 within versus across the two experiments, or combining across both experiments (Fig. S3). In every case
220 we found that across-patient models trained using data from all other patients out-performed within-
221 patient models trained on data only from the subject contributing the given electrode ($ts(XXX) >$

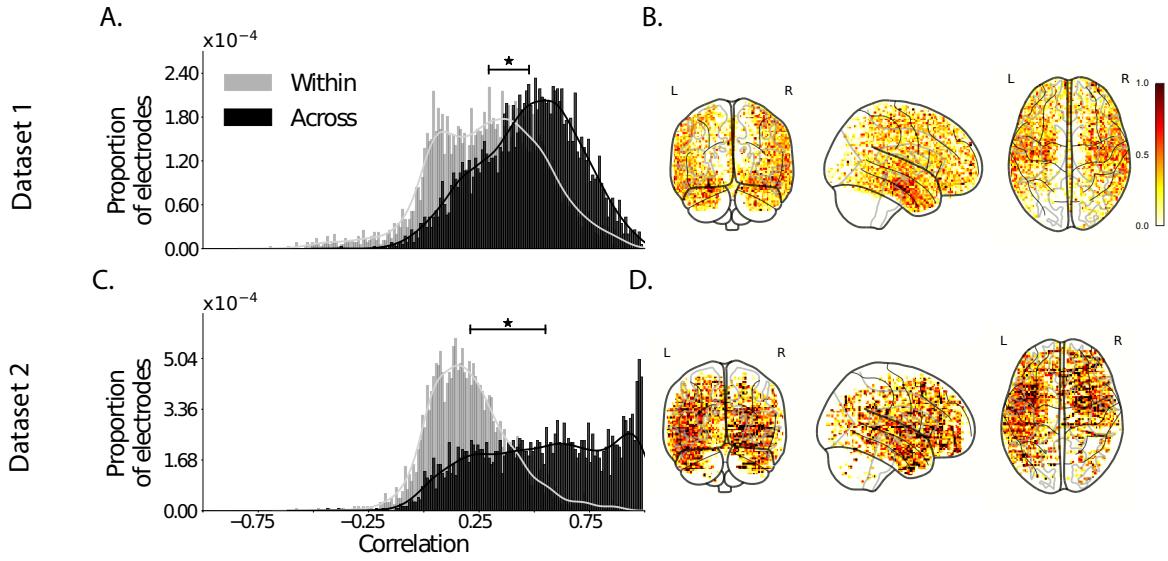


Figure 3: Reconstruction quality across all electrodes in two ECoG datasets. A. Distributions of correlations between observed versus reconstructed activity by electrode, for Dataset 1. The across-patient distribution (black) reflects reconstruction performance using a correlation model trained on all but one patient, and then applied to the held-out patient's data. The within-patient distribution (gray) reflects performance using a correlation model trained on the same patient who contributed the to-be-reconstructed electrode. **B. Distributions of correlations for Dataset 2.** This panel is in the same format as Panel A, but reflects results obtained from Dataset 2. **C.-D. Reconstruction performance by location.** Each dot reflects the location of a single implanted electrode from Dataset 1 (Panel C) or Dataset 2 (Panel D). The dot colors denote the average across-session correlation, using the across-patient correlation model, between the observed and reconstructed activity at the given electrode location.

222 $XXX, ps < XXX$). All reconstruction accuracies also reliably exceeded chance performance ($ts(XXX) >$
223 $XXX, ps < XXX$). Average reconstruction accuracy was highest for the across-patient models limited
224 to data from the same experiment (mean accuracy: 0.55), next-highest for the across-patient models that
225 combined data across both experiments (mean accuracy: 0.50), and lowest for models trained across task
226 (mean accuracy: 0.37). This result also held for each of the Dataset 2 experiments individually (Fig. S4).
227 Taken together, these results indicate that there are reliable commonalities in the spatial correlations of
228 full-brain activity across tasks, but that there are also reliable differences in these spatial correlations
229 across tasks. Whereas reconstruction accuracy benefits from incorporating data from other patients, re-
230 construction accuracy is highest when constrained to within-task data, or data that includes a variety of
231 tasks (e.g., Dataset 1, or combining across the two Dataset 2 experiments).

232 Although both datasets we examined provide good full-brain coverage (when considering data from
233 every patient; e.g. Fig. 3B, D), electrodes are not placed uniformly throughout the brain. For example,
234 electrodes are more likely to be implanted in regions like the medial temporal lobe (MTL), and are rarely
235 implanted in occipital cortex (Fig. 4A, B). Separately for each dataset, for each voxel in the 4 mm^3 voxel
236 MNI152 brain, we computed the proportion of electrodes contained within a 20 MNI unit radius sphere
237 centered on that voxel. We defined the *density* at that location as this proportion. Across Datasets 1
238 and 2, the electrode placement densities were similar (correlation by voxel: $r = 0.57, p < 10^{-10}$). We
239 wondered whether regions with good coverage might be associated with better reconstruction accuracy
240 (e.g. Fig. 3B, D indicate that many electrodes in the MTL have relatively high reconstruction accuracy,
241 and occipital electrodes tend to have relatively low reconstruction accuracy). To test whether this held
242 more generally across the entire brain, for each dataset we computed the electrode placement density for
243 each electrode from each patient (using the proportion of *other* patients' electrodes within 20 MNI units
244 of the given electrode). We then correlated these density values with the across-patient reconstruction
245 accuracies for each electrode. Contrary to our expectation, rather than positive correlations, we found
246 weak (but reliable) negative correlations between reconstruction accuracy and density for both datasets

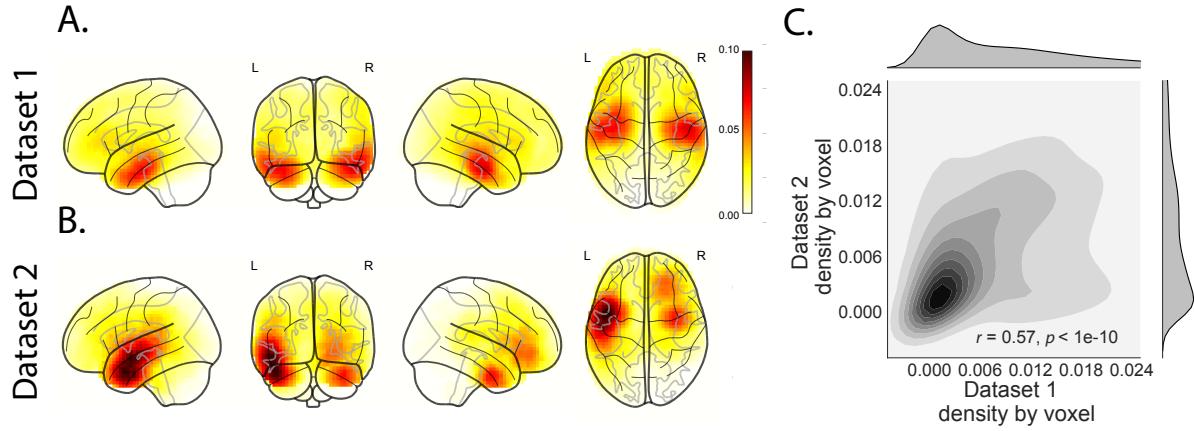


Figure 4: Electrode sampling density by location. **A. Electrode sampling density by voxel in Dataset 1.** Each voxel is colored by the proportion of total electrodes in the dataset that are located within a 20 MNI unit radius sphere centered on the given voxel. **B. Electrode sampling density by voxel in Dataset 2.** This panel displays the sampling density map for Dataset 2, in the same format as Panel A. **C. Correspondence in sampling density by voxel across Datasets 1 and 2.** The two-dimensional histogram displays the by-voxel densities in the two Datasets, and the one-dimensional histograms display the proportions of voxels in each dataset with the given density value. The correlation reported in the panel is across voxels in the 4 mm^3 MNI brain.

(Dataset 1: $r = XXX, p = XXX$; Dataset 2: $r = -0.16, p < 10^{-10}$). This indicates that the reconstruction accuracies we observed are not driven solely by sampling density, but rather may also reflect higher order properties of neural dynamics such as functional correlations between distant voxels [28].

In neurosurgical applications where one wishes to infer full-brain activity patterns, can our framework yield insights into where the electrodes should be placed? A basic assumption of our approach (and of most prior ECoG work) is that electrode recordings are most informative about the neural activity near the recording surface of the electrode. But if we consider that activity patterns throughout the brain are meaningfully correlated, are there particular implantation locations that, if present in a patient's brain, yield especially high reconstruction accuracies throughout the rest of the brain? For example, one might hypothesize that brain structures that are heavily interconnected with many other structures could be more informative about full-brain activity patterns than comparatively isolated structures.

To gain insights into whether particular electrode locations might be especially informative, we first

259 computed the average reconstruction accuracy across all of each patient's electrodes (using the across-
260 patients cross validation test; black histograms in Fig. 3A and C). We labeled each patient's electrodes in
261 each dataset with the average reconstruction accuracy for that patient. In other words, we assigned every
262 electrode from each patient the same value, reflecting how well the activity patterns at those electrodes
263 were reconstructed on average. Next, for each voxel in the 4 mm³ MNI brain, we computed the average
264 value across any electrode (from any patient) that came within 20 MNI units of that voxel's center.
265 Effectively, we computed the average reconstruction accuracy across any patients with electrodes near
266 each voxel, where the averages were weighted to reflect patients who had more electrodes implanted near
267 that location.

268 **JRM STOPPED HERE**

269 To gain insights into this question, we computed the average reconstruction accuracy for each patient,
270 then computed the average patient reconstruction correlation for any patients who had electrodes within a
271 20 MNI unit diameter sphere centered on each voxel location. The resulting maps highlight the locations
272 of implanted electrodes from patients whose reconstructions were especially accurate (Fig. 5A and B).
273 We found that the most informative locations were consistent across datasets which lends support to the
274 notion that different electrode location are more informative about activity across patients (Fig. 5C); $r =$
275 $0.22, p < 10^{-10}$). The locations in dark red might therefore be good candidate implantation targets for
276 neurosurgeons and neurologists who wish to use SuperEEG to reconstruct full-brain electrophysiological
277 signals. The above findings, that one can infer brain activity throughout a person's brain using recordings
278 from a limited number of locations from that person's brain in conjunction with recordings from other
279 people's brains, have deep implications for the structure of brain data. The first implication is that the
280 correlational structure of different people's brain data is largely preserved across individuals. Despite
281 recent evidence that different people have stable but reliably different resting state connectome [29], our
282 results suggest that the correlational structure of different people's brain data is preserved enough across
283 individuals to provide meaningful information.

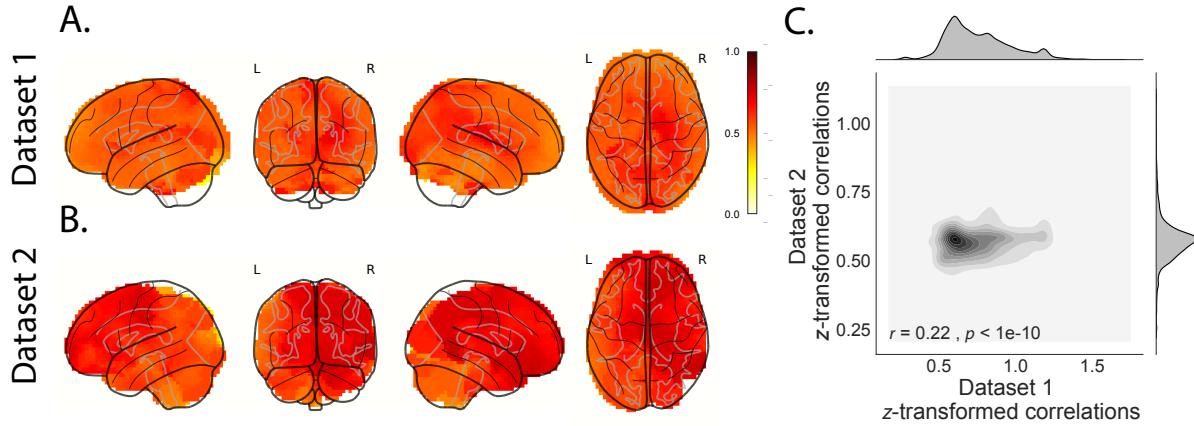


Figure 5: **Most informative electrode locations.** **A. & B.** The glass brain maps displays the average reconstruction correlations (by patient, across all electrodes) for patients with electrodes within a 20 MNI unit diameter sphere centered on each location for dataset 1 and dataset 2. **C.** Correlation between z-transformed correlations by voxel for dataset 1 vs. dataset 2.

284 Discussion

285 SuperEEG infers full-brain activity patterns by leveraging correlations in those patterns of brain activity
 286 within and across people. Although the approach may, in principle, be used to infer brain activity *any-*
 287 *where* in the brain, the inferences perform slightly better for regions with dense electrode sampling across
 288 patients. (Taken to the logical extreme, we could not hope to accurately recover activity patterns from
 289 brain areas where no recordings existed from any patient.) As more data are included in the inference
 290 procedure, this suggests that reconstruction accuracy should improve.

291 A fundamental assumption of the SuperEEG algorithm is that the data covariance matrix is stable
 292 over time and across people. This is a useful simplification. However, a growing body of evidence
 293 from the fMRI community suggests that the data covariance matrix changes in meaningful ways over
 294 time (for example, the data covariance matrix changes from moment-to-moment during story listening,
 295 serving as a unique “fingerprint” for each moment of the story; further, these task-driven timepoint-
 296 specific covariance fingerprints appear to be largely preserved across people [30, 31]). These findings
 297 indicate that the full-brain covariance matrix is not stable over time. Other recent work has shown that

298 people's resting state connectivity matrices may be used to uniquely identify individuals and predict
299 fluid intelligence scores [29]. This indicates that the full-brain covariance matrix is not stable across
300 people. If the fundamental stability assumptions that SuperEEG relies on are violated, how can the
301 SuperEEG algorithm still accurately recover LFP data? It is important to recognize that the fact that
302 variability (over time or across people) is predictive (e.g., of cognitive states during story listening or
303 fluid intelligence scores) does not necessarily mean that this variability is large in magnitude. Rather,
304 we have long known that brain structure is tightly preserved across individuals (and over time, at least
305 on the timescale of typical clinical and experimental recording sessions), and any functional changes
306 must occur within the framework of the underlying structural anatomy. Nevertheless, one could imagine
307 future improvements to the SuperEEG approach that leverage resting state fMRI or structural data [e.g.,
308 diffusion tensor imaging (DTI)] to estimate Bayesian priors over the correlation matrices inferred, in the
309 current framing, using only ECoG data. Further, relaxing the assumption that the covariance matrix is
310 stable (over time and/or across people), and/or incorporating more detailed brain conductance models
311 (e.g., informed by structural MRI scans) may improve the predictive performance of the approach.

312 One potential limitation of the SuperEEG approach is that the above assumption of covariance sta-
313 bility across people may be violated even more if different patients are performing different cognitive
314 tasks. To understand of the extent to which the current findings generalize across cognitive tasks, we
315 replicated our initial findings using a dataset in which patients participated in two tasks, and limited the
316 training data to either within task, across task, or using both tasks. Since we found the most accurate
317 reconstructions using task-specific data, this would suggest building up new databases for estimating
318 each task-specific covariance matrix. Or, using a more sophisticated approach, one could create a hierar-
319 chical model whereby each task-specific covariance matrix was modeled as a perturbation of a "global"
320 task-unspecific covariance matrix (which could in turn be informed by fMRI or DTI data).

321 A second potential limitation of the SuperEEG approach is that it does not provide a natural means
322 of estimating the precise timing of single-neuron action potentials. Prior work has shown that gamma

323 band and broadband activity in the LFP may be used to estimate the firing rates of neurons that un-
324 derly the population contributing to the LFP [32]. Because SuperEEG reconstructs LFPs throughout the
325 brain, one could in principle use gamma or broadband power in the reconstructed signals to estimate the
326 corresponding firing rates (though not the timings of individual action potentials).

327 Beyond providing a means of estimating ongoing activity throughout the brain using already im-
328 planted electrodes, our work also has implications for where to place the electrodes in the first place.
329 Electrodes are typically implanted to maximize coverage of suspected epileptogenic tissue. However,
330 our findings suggest that this approach could be further optimized. Specifically, one could leverage not
331 only the non-invasive recordings taken during an initial monitoring period (as is currently done), but also
332 recordings collected from other patients. We could then ask: given everything we know about the other
333 patients and from the scalp recordings of this new patient, where should we place a fixed number of
334 electrodes to maximize our ability to map seizure foci? As shown in Figure 5, recordings from different
335 locations are differently informative in terms of reconstructing the spatiotemporal patterns throughout
336 the brain. This property might be leveraged in decisions about where to surgically implant electrodes in
337 future patients.

338 **Concluding remarks**

339 Over the past several decades, neuroscientists have begun to leverage the strikingly profound mathemati-
340 cal structure underlying the brain’s complexity to infer how our brains carry out computations to support
341 our thoughts, actions, and physiological processes. Whereas traditional beamforming techniques rely on
342 geometric source-localization of signals measured at the scalp, here we propose an alternative approach
343 that leverages the rich correlational structure of a large dataset of human intracranial recordings. In do-
344 ing so, we are one step closer to observing, and perhaps someday understanding, the full spatiotemporal
345 structure of human neural activity.

³⁴⁶ **Code availability**

³⁴⁷ We have released an open-source SuperEEG Python toolbox. All of the code used in this manuscript is
³⁴⁸ on GitHub, and the code may be shared using a GitHub account accessible to the reviewers upon request.

³⁴⁹ **Data availability**

³⁵⁰ The dataset analyzed in this study was generously shared by Michael J. Kahana. A portion of the dataset
³⁵¹ may be downloaded [here](#).

³⁵² **Acknowledgements**

³⁵³ We are grateful for useful discussions with Luke J. Chang and Matthijs van der Meer. We are also grateful
³⁵⁴ to Michael J. Kahana for generously sharing the ECoG dataset we analyzed in our paper, which was
³⁵⁵ collected under NIMH grant MH55687 to MJK. Our work was also supported in part by NSF EPSCoR
³⁵⁶ Award Number 1632738. The content is solely the responsibility of the authors and does not necessarily
³⁵⁷ represent the official views of our supporting organizations.

³⁵⁸ **Author Contributions**

³⁵⁹ J.R.M conceived and initiated the project. L.L.W.O. and A.C.H. performed the analyses. J.R.M. and
³⁶⁰ L.L.W.O. wrote the manuscript.

³⁶¹ **Author Information**

³⁶² Reprints and permissions information is available at [www.nature.com/reprints](#). The authors declare no
³⁶³ competing financial interests. Readers are welcome to comment on the online version of the paper.
³⁶⁴ Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps
³⁶⁵ and institutional affiliations. Correspondence and requests for materials should be addressed to J.R.M.
³⁶⁶ (jeremy.r.manning@dartmouth.edu).

367 **References and Notes**

- 368 [1] R. J. Sawyer, *The Terminal Experiment* (HarperPrism, 1995).
- 369 [2] T. J. Sejnowski, P. S. Churchland, J. A. Movshon, *Nature Neuroscience* **17**, 1440 (2014).
- 370 [3] J. Sarvas, *Phys. Med. Biol.* **32**, 11 (1987).
- 371 [4] A. Z. Snyder, *Electroencephalography and Clinical Neurophysiology* **80**, 321 (1991).
- 372 [5] S. Baillet, J. C. Mosher, R. M. Leahy, *IEEE Signal Processing Magazine* **18**, 14 (2001).
- 373 [6] A. Hillebrand, K. D. Singh, I. E. Holliday, P. L. Furlong, G. R. Barnes, *Human Brain Mapping* **25**,
374 199 (2005).
- 375 [7] R. J. Huster, S. Debener, T. Eichele, C. S. Herrmann, *The Journal of Neuroscience* **32**, 6053 (2012).
- 376 [8] C. E. Rasmussen, *Gaussian processes for machine learning* (MIT Press, 2006).
- 377 [9] P. B. Sederberg, M. J. Kahana, M. W. Howard, E. J. Donner, J. R. Madsen, *Journal of Neuroscience*
378 **23**, 10809 (2003).
- 379 [10] P. B. Sederberg, *et al.*, *Cerebral Cortex* **17**, 1190 (2007).
- 380 [11] P. B. Sederberg, *et al.*, *Psychological Science* **18**, 927 (2007).
- 381 [12] J. R. Manning, S. M. Polyn, G. Baltuch, B. Litt, M. J. Kahana, *Proceedings of the National Academy
382 of Sciences, USA* **108**, 12893 (2011).
- 383 [13] J. R. Manning, M. R. Sperling, A. Sharan, E. A. Rosenberg, M. J. Kahana, *The Journal of Neuro-
384 science* **32**, 8871 (2012).
- 385 [14] Y. Ezzyat, *et al.*, *Current Biology* **27**, 1 (2017).
- 386 [15] P. C. Horak, *et al.*, *Epilepsia* **58**, 373 (2017).

- 387 [16] J. E. Kragel, *et al.*, *NeuroImage* **155**, 70 (2017).
- 388 [17] M. T. Kucewicz, *et al.*, *Brain* **140**, 1337 (2017).
- 389 [18] J.-J. Lin, *et al.*, *Hippocampus* **27**, 1040 (2017).
- 390 [19] E. A. Solomon, *et al.*, *Nature Communications* **In press** (2018).
- 391 [20] C. T. Weidemann, *et al.*, *Journal of Experimental Psychology: General* **In press** (2018).
- 392 [21] Y. Ezzyat, *et al.*, *Nature Communications* **9**, <https://www.doi.org/10.1038/s41467> (2018).
- 393 [22] M. T. Kucewicz, *et al.*, *Brain* **141**, 971 (2018).
- 394 [23] B. T. T. Yeo, *et al.*, *Journal of Neurophysiology* **106**, 1125 (2011).
- 395 [24] G. Grabner, *et al.*, *Medical Image Computing and Computer-Assistend Intervention* **9**, 58 (2006).
- 396 [25] J. R. Manning, R. Ranganath, K. A. Norman, D. M. Blei, *PLoS One* **9**, e94914 (2014).
- 397 [26] J. R. Manning, *et al.*, *NeuroImage* (2018).
- 398 [27] C. O. Becker, *et al.*, *Scientific Reports* **8**, <https://www.doi.org/10.1038/s41598> (2018).
- 399 [28] R. F. Betzel, *et al.*, *arXiv* p. 1706.06088 (2017).
- 400 [29] E. S. Finn, *et al.*, *Nature Neuroscience* **18**, 1664 (2015).
- 401 [30] E. Simony, C. J. Honey, J. Chen, U. Hasson, *Nature Communications* **7**, 1 (2016).
- 402 [31] J. R. Manning, *et al.*, *bioRxiv* p. 106690 (2017).
- 403 [32] J. R. Manning, J. Jacobs, I. Fried, M. J. Kahana, *The Journal of Neuroscience* **29**, 13613 (2009).