

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

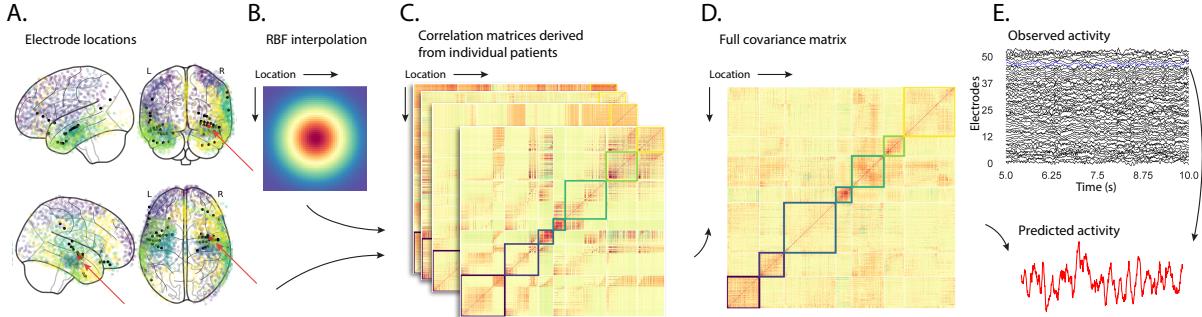


Figure 1: Methods overview. **A. Electrode locations.** Each dot reflects the location of a single electrode in dataset 1, colored according to 7 factor labels (see Panel D for details). One patient’s electrode locations are highlighted in black and the to-be-reconstructed recording location is highlighted in red. **B. Radial basis function (RBF).** 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. Combined correlation matrix.** We estimate a single full-brain correlation matrix by averaging the patient-specific correlation matrices. We sort the resulting correlation matrix based on 7 factor labels obtained from k-means clustering [23]. **E. Reconstructing activity throughout the brain.** Given the observed activity from the patient’s electrodes and the estimated correlation matrix (Panel D), we can compute a maximum likelihood estimate of the voltage trace at any location in \bar{R} . An example reconstruction (at the red dot in Panel A) is shown in red, and the actual recording at that location is highlighted above in blue.

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 every

recording (from every electrode). We then excluded any electrodes that showed putative epileptiform activity. Specifically, we excluded from further analysis any electrode that exhibited an average kurtosis of 10 or greater across all of that patient’s recording sessions. We also excluded any patients with fewer than 2 electrodes that passed this criteria, as the SuperEEG algorithm requires measuring correlations between 2 or more electrodes from each patient. For Dataset 1, this yielded clean recordings from 4168 electrodes implanted throughout the brains of 67 patients (Fig. 1A); for Dataset 2, this yielded clean recordings from 2975 electrodes from 24 patients. Each individual patient contributes electrodes from a limited set of brain locations, which we localized in a common space [MNI152; 24]; an example Dataset 1 patient’s 54 electrodes that passed the predefined 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. 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

98 work as a guide [25, 26], we set $\lambda = 20$, although this could in theory be optimized, e.g., using cross
 99 validation or a formal model [e.g., 26].

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

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

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

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

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

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

112 Given this weight matrix, W_s , and the observed inter-electrode correlation matrix for patient s , \bar{C}_s ,
 113 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):

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

Intuitively, the numerators capture the general structures of the patient-specific estimates of full-brain correlations, and the denominators account for which locations were near the implanted electrodes in each patient. To obtain \hat{K} , we compute a weighted average across the estimated patient-specific full-brain correlation matrices, where patients with observed electrodes near a particular set of locations in \hat{K} contribute more to the estimate.

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

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 of all other locations in \bar{R} . In other words, β_s reflects the locations in \bar{R} where we did not observe a recording for patient s (these are the recording locations we will want to fill in using SuperEEG). We can sub-divide \hat{K} as follows:

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

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

¹²⁶ Here $\hat{K}_{\beta_s, \alpha_s}$ represents the correlations between the “unknown” activity at the locations in β_s and the
¹²⁷ observed activity at the locations in α_s , and $\hat{K}_{\alpha_s, \alpha_s}$ represents the correlations between the observed
¹²⁸ recordings (at the locations in α_s).

¹²⁹ Let Y_{s,k,α_s} be the number-of-timepoints (T) by $\text{length}(\alpha_s)$ matrix of (observed) voltages from the
¹³⁰ electrodes in α_s during session k from patient s . Then we can estimate the voltage from patient s ’s k^{th}
¹³¹ session at the locations in β_s using [8]:

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

¹³² This equation is the foundation of the SuperEEG algorithm. Whereas we observe recordings only at the
¹³³ locations in α_s , Equation 12 allows us to estimate the recordings at all locations in β_s , which we can
¹³⁴ define *a priori* to include any locations we wish, throughout the brain. This yields estimates of the time-
¹³⁵ varying voltages at *every* location in \bar{R} , provided that we define \bar{R} in advance to include the union of all
¹³⁶ of the locations in α_s and all of the locations at which we wish to estimate recordings (i.e., a timeseries
¹³⁷ of voltages).

¹³⁸ We designed our approach to be agnostic to electrode impedances, as electrodes that do not exist
¹³⁹ do not have impedances. Therefore our algorithm recovers voltages in standard deviation (z -scored)
¹⁴⁰ units rather than attempting to recover absolute voltages. (This property reflects the fact that $\hat{K}_{\beta_s, \alpha_s}$ and
¹⁴¹ $\hat{K}_{\alpha_s, \alpha_s}$ are correlation matrices rather than covariance matrices.) Also, note that Equation 12 directly
¹⁴² requires computing a T by T matrix, which can become computationally intractable when T is very
¹⁴³ large (e.g., for the patient highlighted in Fig. 2, $T = 20458799$). However, because Equation 12 is time
¹⁴⁴ 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
¹⁴⁵ the corresponding samples from Y_{s,k,α_s}).

¹⁴⁶ The SuperEEG algorithm described above and in Figure 1 allows us to estimate (up to a constant
¹⁴⁷ scaling factor) LFPs for each patient at all arbitrarily chosen locations in the set \bar{R} , *even if we did not*

¹⁴⁸ record that patient's brain at all of those locations. We next turn to an evaluation of the accuracy of those
¹⁴⁹ estimates.

¹⁵⁰ **JRM STOPPED HERE**

¹⁵¹ **Results**

¹⁵² To test the accuracy with which the SuperEEG algorithm reconstructs activity throughout the brain, we
¹⁵³ held out each electrode from the full dataset in turn and treated it as unobserved. We then asked: how
¹⁵⁴ closely did each of the SuperEEG-reconstructed LFPs match the observed data? We sought to evaluate
¹⁵⁵ both the overall reconstruction accuracy as well as how reconstruction accuracy varied as a function of
¹⁵⁶ implantation location.

¹⁵⁷ We first examined raw LFP traces and their associated Super EEG-derived reconstructions. Figure 2A
¹⁵⁸ displays the LFP from the red electrode in Figure 1A, and its associated reconstruction, during a 5 s time
¹⁵⁹ window during one of the patient's 6 recording sessions. Figure 2B displays a 2D histogram of the
¹⁶⁰ observed versus reconstructed voltages for every sample across 14.2 total hours of recordings from that
¹⁶¹ patient (correlation: $r = 0.91, p < 10^{-10}$). Although the SuperEEG algorithm recovered the recordings
¹⁶² from this electrode well, we sought to quantify the algorithm's performance across the full dataset.

¹⁶³ Holding out each electrode from each patient in turn, we computed the average correlation (across
¹⁶⁴ recording sessions) between the Super EEG-reconstructed voltage traces and the observed voltage traces
¹⁶⁵ from that electrode. For each reconstruction, we estimated the full-brain correlation matrix using every
¹⁶⁶ other patient's data (i.e., every patient except the one who contributed the to-be-reconstructed electrode
¹⁶⁷ data). In our analyses, we then substituted the average correlation matrix computed after excluding
¹⁶⁸ patient s 's data for \hat{K} in Equations 10 and 11. This step ensured that the data we were reconstructing
¹⁶⁹ could not also be used to estimate the between-location correlations that drove the reconstructions via
¹⁷⁰ Equation 12 (otherwise the analysis would be circular).

¹⁷¹ We obtained a single correlation coefficient for each electrode location in \bar{R} , reflecting how well

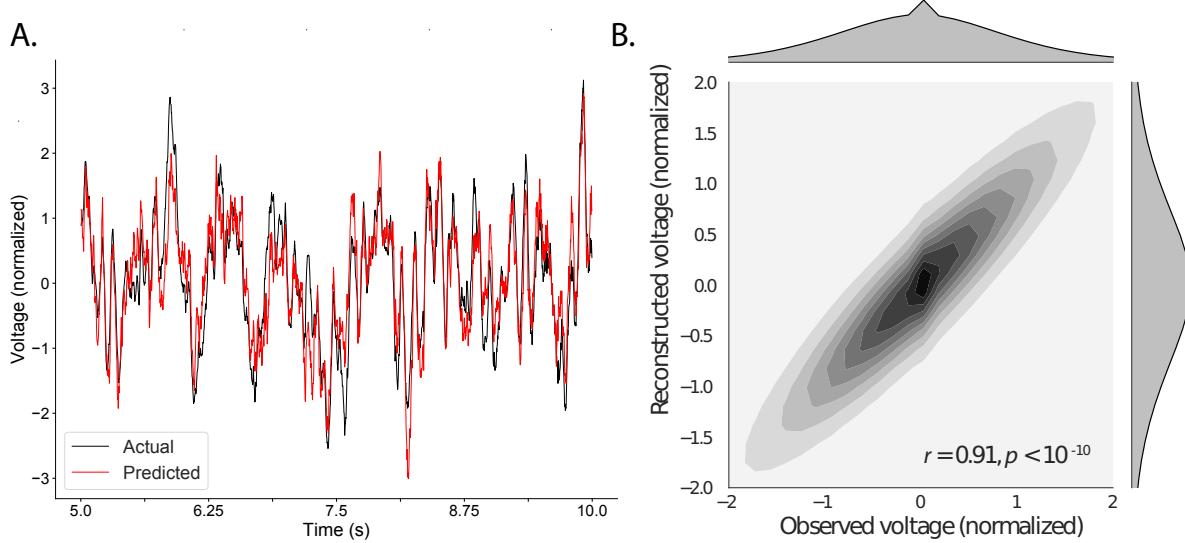


Figure 2: Observed and reconstructed LFP from a single electrode. A. Example LFP. A 2 s recording from the blue electrode in Figure 1A is displayed in red, and the reconstructed LFP during the same time window is shown in black. All voltages are plotted in standard deviation units. **B. Observed versus reconstructed voltages over 14.2 hours.** The 2D histogram reflects the relation between distributions of observed versus reconstructed voltages from one patient, across the 14.2 hours of recorded data collected in 6 recording sessions. The correlation reported in the panel is between the observed and reconstructed voltages.

172 the SuperEEG algorithm was able to recover the recording at that location by incorporating data across
173 patients (Across shown in black, see Fig. 3A). We also reconstructed activity for each electrode using a
174 model trained on the remaining electrodes from only that patient, to account for reconstruction accuracy
175 attributed to volume conductance alone (Within shown in gray, see Fig. 3A). For the first dataset, we
176 compared these two distributions of correlation coefficients (paired *t*-test between *z*-transformed mean
177 correlation coefficients by patient: $t(66) = 9.64, p < 10^{-10}$). We repeated this analysis on a similar
178 dataset (Fig. 3C) with similar results (paired *t*-test between *z*-transformed mean correlation coefficients
179 by patient: $t(23) = 6.93, p < 10^{-5}$). This is an especially conservative test, given that the SuperEEG
180 reconstructions exclude (from the correlation matrix estimates) all data from the patient whose data is
181 being reconstructed. Furthermore, we also replicated this finding for each independent experiment within
182 dataset 2 (Fig. S3 (paired *t*-test between *z*-transformed mean correlation coefficients by patient for exper-
183 iment 1: $t(23) = 6.23, p < 10^{-5}$ and experiment 2: $t(23) = 6.62, p < 10^{-5}$). That the SuperEEG-derived
184 correlations were reliably stronger than these correlations obtained using a volume conductance null
185 model is exciting for two reasons. First, it implies that distant electrodes provide additional predictive
186 power to the data reconstructions beyond the information contained in nearby electrodes. Second, it
187 implies that the spatial correlations driving the SuperEEG algorithm are, to some extent, shared across
188 people.

189 We were interested in the task specific contributions to the reconstruction accuracy. Each patient
190 in the second dataset participated in two free recall experiments. We ran similar analyses for both
191 experiments and found that activity was best reconstructed when limiting the training data to within
192 task, as opposed to across task or incorporating data from both tasks (Fig. S1 (mean reconstruction
193 accuracy incorporating data within task: 0.55, across task: 0.37, all tasks: .50)). Although reconstruction
194 accuracy in the across task analysis was still better than the volume conductance model alone (paired
195 *t*-test between *z*-transformed mean correlation coefficients by patient: $t(47) = 5.65, p < 10^{-5}$), these
196 results suggests that having a common tasks for patients may yield better reconstruction accuracy.

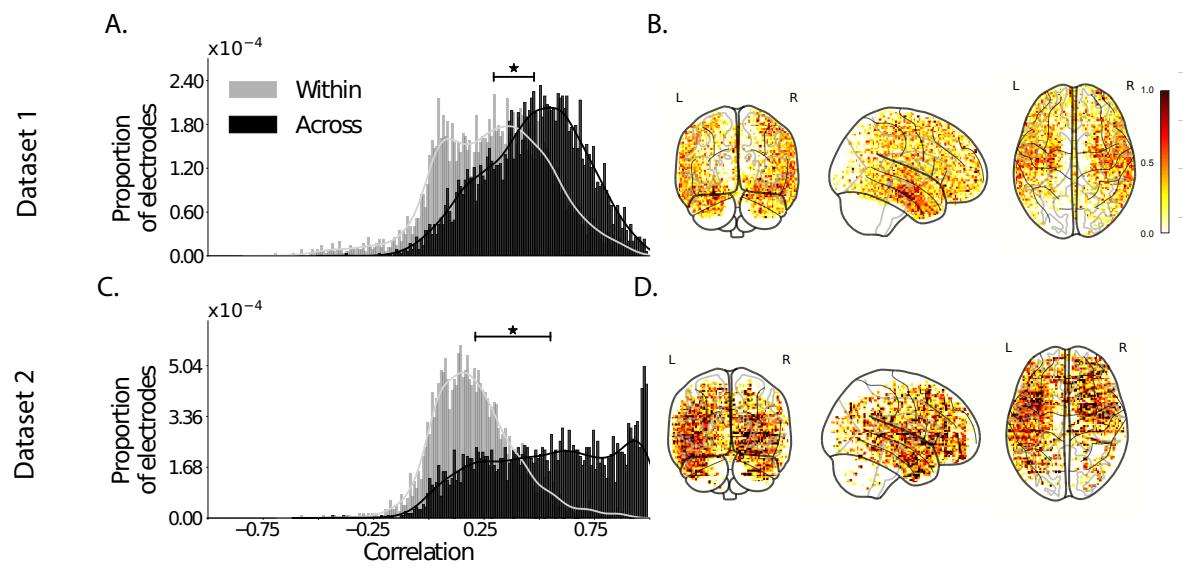


Figure 3: Reconstruction quality. A. & C. Distributions of correlation coefficients. Across all electrodes from all patients in the labeled dataset, the panel displays the distribution of correlations between the observed and reconstructed LFP data using models trained on data from all other patients (Across, in black) and all other electrodes from the same patient (Within, in gray). **B. & D. Correlation maps.** The glass brain maps display the average correlation between the observed LFP data and the across-subjects model reconstructed data by location, for each labeled experiment.

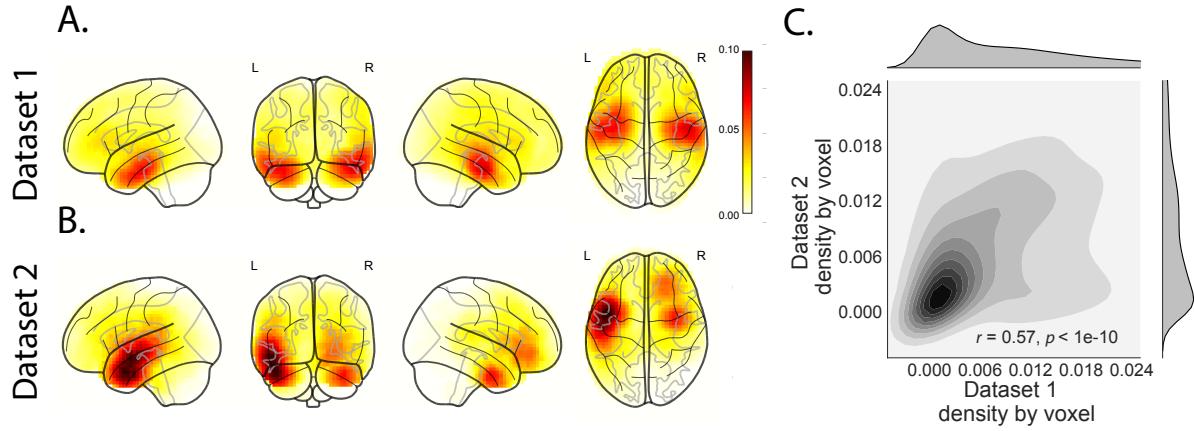


Figure 4: Sampling density and reconstruction quality. **A. & B.** The glass brain maps show sampling density by voxel location for dataset 1 and dataset 2. **C.** Correlation of sampling density by voxel location for dataset 1 vs. dataset 2.

197 We also wondered whether reconstruction quality (measured as the correlation between the observed
 198 and reconstructed data) varied with the electrode locations (Fig. 3B & D). In general, reconstruction
 199 quality remained high throughout the brain. Although reconstruction accuracy appeared high in the me-
 200 dial temporal lobe, which is a common epileptic focus (and therefore a common target for electrode
 201 implantation), we observed a weak but statistically reliable negative correlation between reconstruction
 202 quality and electrode density (defined as the proportion of electrodes within 20 MNI units for each loca-
 203 tion; dataset 1: $r = -0.07, p < 10^{-5}$, dataset 2: $r = -0.16, p < 10^{-10}$). This provides some evidence
 204 that our reconstruction accuracy results cannot be driven only by volume conductance. Qualitatively, it
 205 appeared that the distribution of electrodes was similar across the datasets, suggesting potential com-
 206 monalities of target locations across patients and similarities in surgical decisions. Indeed, we found
 207 a relatively strong correlation between the electrode densities within the two datasets (defined as the
 208 proportion of electrodes within 20 MNI units for each 34686 voxels (Fig. 4A, B); $r = 0.57, p < 10^{-10}$).

209 In addition to exploring how reconstruction quality varies with location, we also wondered whether
 210 there might be effects of electrode placements on reconstruction quality. For example, are there particular
 211 implantation locations that yield especially high reconstruction accuracies at other locations throughout

the brain? To gain insights into this questions, we computed the average reconstruction correlation for each patient, then computed the average patient reconstruction correlation for any patients who had electrodes within a 20 MNI unit diameter sphere centered on each voxel location. The resulting maps highlight the locations of implanted electrodes from patients whose reconstructions were especially accurate (Fig. 5A and B). We found that the most informative locations were consistent across datasets which lends support to the notion that different electrode location are more informative about activity across patients (Fig. 5C); $r = 0.22, p < 10^{-10}$). The locations in dark red might therefore be good candidate implantation targets for neurosurgeons and neurologists who wish to use SuperEEG to reconstruct full-brain electrophysiological signals. The above findings, that one can infer brain activity throughout a person’s brain using recordings from a limited number of locations from that person’s brain in conjunction with recordings from other people’s brains, have deep implications for the structure of brain data. The first implication is that the correlational structure of different people’s brain data is largely preserved across individuals. Despite recent evidence that different people have stable but reliably different resting state connectome [28], our results suggest that the correlational structure of different people’s brain data is preserved enough across individuals to provide meaningful information.

Discussion

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

A fundamental assumption of the SuperEEG algorithm is that the data covariance matrix is stable over time and across people. This is a useful simplification. However, a growing body of evidence

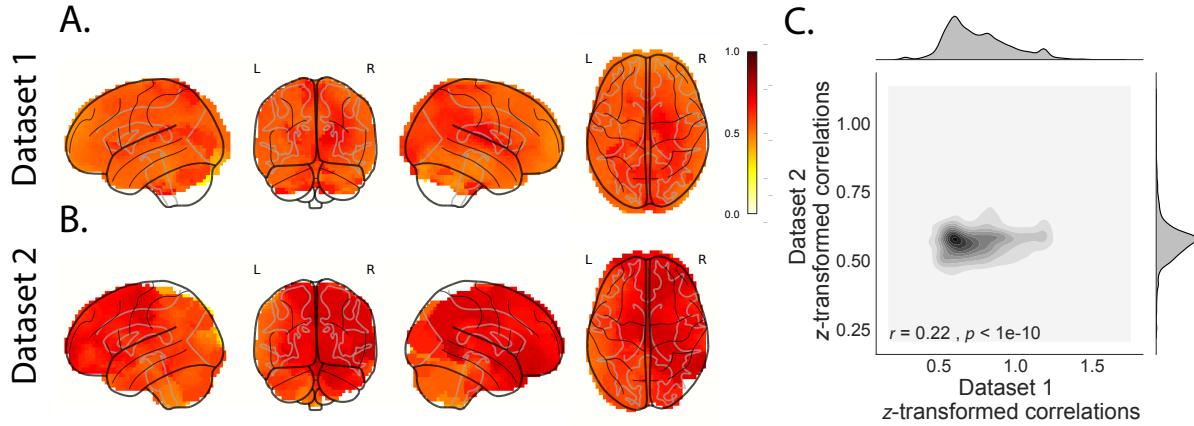


Figure 5: Most informative electrode locations. **A. & B.** The glass brain maps displays the average reconstruction correlations (by patient, across all electrodes) for patients with electrodes within a 20 MNI unit diameter sphere centered on each location for dataset 1 and dataset 2. **C.** Correlation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

from the fMRI community suggests that the data covariance matrix changes in meaningful ways over time (for example, the data covariance matrix changes from moment-to-moment during story listening, serving as a unique “fingerprint” for each moment of the story; further, these task-driven timepoint-specific covariance fingerprints appear to be largely preserved across people [29, 30]). These findings indicate that the full-brain covariance matrix is not stable over time. Other recent work has shown that people’s resting state connectivity matrices may be used to uniquely identify individuals and predict fluid intelligence scores [28]. This indicates that the full-brain covariance matrix is not stable across people. If the fundamental stability assumptions that SuperEEG relies on are violated, how can the SuperEEG algorithm still accurately recover LFP data? It is important to recognize that the fact that variability (over time or across people) is predictive (e.g., of cognitive states during story listening or fluid intelligence scores) does not necessarily mean that this variability is large in magnitude. Rather, we have long known that brain structure is tightly preserved across individuals (and over time, at least on the timescale of typical clinical and experimental recording sessions), and any functional changes must occur within the framework of the underlying structural anatomy. Nevertheless, one could imagine future improvements to the SuperEEG approach that leverage resting state fMRI or structural data [e.g.,

251 diffusion tensor imaging (DTI)] to estimate Bayesian priors over the correlation matrices inferred, in the
252 current framing, using only ECoG data. Further, relaxing the assumption that the covariance matrix is
253 stable (over time and/or across people), and/or incorporating more detailed brain conductance models
254 (e.g., informed by structural MRI scans) may improve the predictive performance of the approach.

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

264 A second potential limitation of the SuperEEG approach is that it does not provide a natural means
265 of estimating the precise timing of single-neuron action potentials. Prior work has shown that gamma
266 band and broadband activity in the LFP may be used to estimate the firing rates of neurons that un-
267 derly the population contributing to the LFP [31]. Because SuperEEG reconstructs LFPs throughout the
268 brain, one could in principle use gamma or broadband power in the reconstructed signals to estimate the
269 corresponding firing rates (though not the timings of individual action potentials).

270 Beyond providing a means of estimating ongoing activity throughout the brain using already im-
271 planted electrodes, our work also has implications for where to place the electrodes in the first place.
272 Electrodes are typically implanted to maximize coverage of suspected epileptogenic tissue. However,
273 our findings suggest that this approach could be further optimized. Specifically, one could leverage not
274 only the non-invasive recordings taken during an initial monitoring period (as is currently done), but also
275 recordings collected from other patients. We could then ask: given everything we know about the other

276 patients and from the scalp recordings of this new patient, where should we place a fixed number of
277 electrodes to maximize our ability to map seizure foci? As shown in Figure 5, recordings from different
278 locations are differently informative in terms of reconstructing the spatiotemporal patterns throughout
279 the brain. This property might be leveraged in decisions about where to surgically implant electrodes in
280 future patients.

281 **Concluding remarks**

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

289 **Code availability**

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

292 **Data availability**

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

295 **Acknowledgements**

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

301 **Author Contributions**

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

304 **Author Information**

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

310 **References and Notes**

- 311 [1] R. J. Sawyer, *The Terminal Experiment* (HarperPrism, 1995).
- 312 [2] T. J. Sejnowski, P. S. Churchland, J. A. Movshon, *Nature Neuroscience* **17**, 1440 (2014).
- 313 [3] J. Sarvas, *Phys. Med. Biol.* **32**, 11 (1987).
- 314 [4] A. Z. Snyder, *Electroencephalography and Clinical Neurophysiology* **80**, 321 (1991).

- 315 [5] S. Baillet, J. C. Mosher, R. M. Leahy, *IEEE Signal Processing Magazine* **18**, 14 (2001).
- 316 [6] A. Hillebrand, K. D. Singh, I. E. Holliday, P. L. Furlong, G. R. Barnes, *Human Brain Mapping* **25**,
- 317 199 (2005).
- 318 [7] R. J. Huster, S. Debener, T. Eichele, C. S. Herrmann, *The Journal of Neuroscience* **32**, 6053 (2012).
- 319 [8] C. E. Rasmussen, *Gaussian processes for machine learning* (MIT Press, 2006).
- 320 [9] P. B. Sederberg, M. J. Kahana, M. W. Howard, E. J. Donner, J. R. Madsen, *Journal of Neuroscience*
- 321 **23**, 10809 (2003).
- 322 [10] P. B. Sederberg, *et al.*, *Cerebral Cortex* **17**, 1190 (2007).
- 323 [11] P. B. Sederberg, *et al.*, *Psychological Science* **18**, 927 (2007).
- 324 [12] J. R. Manning, S. M. Polyn, G. Baltuch, B. Litt, M. J. Kahana, *Proceedings of the National Academy*
- 325 *of Sciences, USA* **108**, 12893 (2011).
- 326 [13] J. R. Manning, M. R. Sperling, A. Sharan, E. A. Rosenberg, M. J. Kahana, *The Journal of Neuro-*
- 327 *science* **32**, 8871 (2012).
- 328 [14] Y. Ezzyat, *et al.*, *Current Biology* **27**, 1 (2017).
- 329 [15] P. C. Horak, *et al.*, *Epilepsia* **58**, 373 (2017).
- 330 [16] J. E. Kragel, *et al.*, *NeuroImage* **155**, 70 (2017).
- 331 [17] M. T. Kucewicz, *et al.*, *Brain* **140**, 1337 (2017).
- 332 [18] J.-J. Lin, *et al.*, *Hippocampus* **27**, 1040 (2017).
- 333 [19] E. A. Solomon, *et al.*, *Nature Communications* **In press** (2018).
- 334 [20] C. T. Weidemann, *et al.*, *Journal of Experimental Psychology: General* **In press** (2018).

- 335 [21] Y. Ezzyat, *et al.*, *Nature Communications* **9**, <https://www.doi.org/10.1038/s41467> (2018).
- 336 [22] M. T. Kucewicz, *et al.*, *Brain* **141**, 971 (2018).
- 337 [23] B. T. T. Yeo, *et al.*, *Journal of Neurophysiology* **106**, 1125 (2011).
- 338 [24] G. Grabner, *et al.*, *Medical Image Computing and Computer-Assisted Intervention* **9**, 58 (2006).
- 339 [25] J. R. Manning, R. Ranganath, K. A. Norman, D. M. Blei, *PLoS One* **9**, e94914 (2014).
- 340 [26] J. R. Manning, *et al.*, *NeuroImage* (2018).
- 341 [27] C. O. Becker, *et al.*, *Scientific Reports* **8**, <https://www.doi.org/10.1038/s41598> (2018).
- 342 [28] E. S. Finn, *et al.*, *Nature Neuroscience* **18**, 1664 (2015).
- 343 [29] E. Simony, C. J. Honey, J. Chen, U. Hasson, *Nature Communications* **7**, 1 (2016).
- 344 [30] J. R. Manning, *et al.*, *bioRxiv* p. 106690 (2017).
- 345 [31] J. R. Manning, J. Jacobs, I. Fried, M. J. Kahana, *The Journal of Neuroscience* **29**, 13613 (2009).