University of Washington CSE Technical Report 2006-06-01

Learning to Predict Cortical Potentials using Simultaneous Transcranial Recording

Aaron P. Shon[1]† **Kai J. Miller**[2]‡ **Jeffrey G. Ojemann** ∗
Mark D. Holmes ∗∗ **Rajesh P. N. Rao**[3] †
†Dept. of Computer Science and Engineering ‡Dept. of Physics
∗ Dept. of Neurosurgery ∗∗ Dept. of Neurology
University of Washington Seattle, WA 98195
{*aaron*[1],*kai*[2],*rao*[3]}*@cs.washington.edu*

Abstract

Characterizing the transform from cortical potentials to scalp electrode readings (the *forward problem* of brain tomography) and from scalp electrode readings to cortical potentials (the *inverse problem*) remain challenging research problems. A major reason for this is the lack of ground truth data on both sides of the mapping: to our knowledge, no published studies to date have analyzed channel characteristics derived from simultaneous high-density cortical and scalp recordings. Using simultaneous recordings from a high-density array of 258 EEG electrodes and 42 subdural ECoG electrodes implanted in a human subject, we present the first such characterization of the cortex-scalp transfer function using known ground truth data. We show that a linear method suffices to solve the forward problem. We apply nonlinear statistical machine learning methods based on non-parametric inference for the inverse problem. Specifically, we show that a Gaussian process method employing minimal assumptions about the transfer function can recover important temporal and spectral features of the underlying brain signal. Applying a particle filter to the cortical estimate improves the model's accuracy in both the temporal and spectral domains.

1 Introduction

Electroencephalography (EEG) and electrocorticography (ECoG) are attractive for a variety of applications, including brain mapping [3, 4, 5, 6] and user interfaces for locked-in patients [12, 20]. The strengths and weaknesses of each approach are well known; to summarize, EEG is non-invasive but has low signal-to-noise ratio, while ECoG is clinically invasive but provides much cleaner signals. If EEG contains meaningful information about cortical activity, it is reasonable to hope that sufficiently expressive, predictive models could reduce the need for invasive surgical procedures. While many physiologically reasonable models have been proposed, the lack of *in vivo*, ground truth data has made verification of these models and associated parameters problematic. Below, we discern between the *forward problem* of determining EEG signals given ECoG signals, and the *inverse problem* of determining ECoG signals given EEG signals.

We present what is, to our knowledge, the first characterization of EEG-ECoG transfer functions based on simultaneous recordings from dense electrode arrays on the scalp and the cortical surface of an awake, interacting human. We begin by analyzing mutual information between individual EEG and ECoG channels. We then characterize the frequency

attenuation of the ECoG → EEG transfer function. Preliminary approaches to solving the forward and inverse regression problems follow. The regression models rely on minimal assumptions. We discuss the performance of our general models, and consider future extensions to the inverse problem using simultaneous recording.

2 Methods

We simultaneously recorded EEG (using 258 electrode leads) and ECoG (using 42 cortical electrodes) from a single epileptic subject. Fig. 1(a) is a diagram of electrode grid placement on the surface of the scalp; (b) shows a sagittal X-ray of ECoG electrode placement. Our analysis is based on 17 trials (1 training, 16 testing) of simultaneous recording while the subject performed a hand-related motor task.

The experiment was performed in the patient's hospital room. Simultaneous recordings were performed using a 258 high density sponge-electrode EGI EEG array and a 42 electrode EcoG array recorded on Neuroscan Synamps2 amplifiers [1]. Data was recorded at 1000 Hz on each amplifier, and was synchronized by aligning a trigger sent simultaneously to each system. Instrument-imposed band-pass filtering was performed from 0.15 to 200 Hz in the ECoG recording and 0.1 to 500 Hz in the EEG recording. The patient was instructed to move their hand when given a visual stimulus to do so; there were thirty 3 second movement stimuli interleaved with 3 second rest periods. Two electrodes from each array were used as reference and ground, and two of the ECoG electrodes were rejected due to defect; our analysis is based on data from the remaining 38 ECoG and 256 EEG electrodes. We performed a spatial common-average reference across the sets of ECoG and EEG electrodes. To minimize the number of model assumptions about the data, we did not perform any additional artifact rejection for the results shown here.

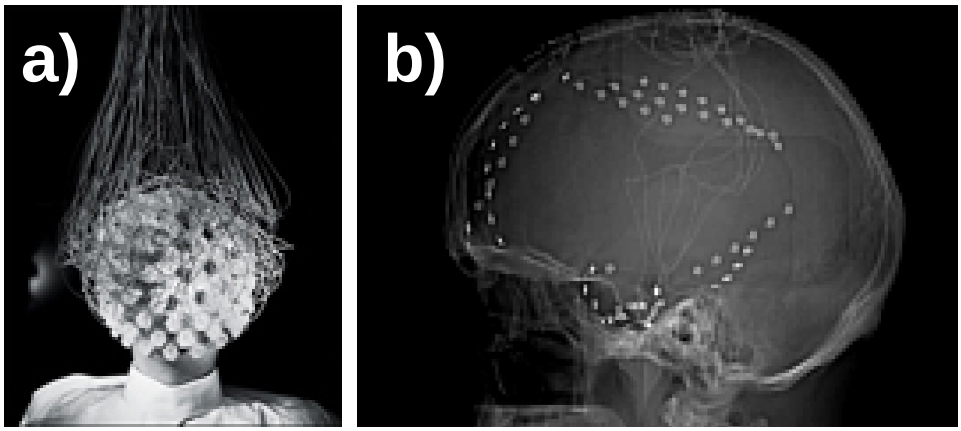

Figure 1: **Simultaneous recording of EEG and ECoG:** (a) Example of EEG electrode placement. (b) Sagittal X-ray of our subject's ECoG electrode placement.

3 Channel characterization

3.1 Maximum entropy feature selection

The high dimensionality of EEG and ECoG data (256 and 38 dimensions respectively) and large number of data points collected ($\sim$44000) motivate feature selection to determine which EEG and ECoG electrodes are most informative. We therefore computed mutual information between pairs of ECoG and EEG electrodes as follows. First, we discretized the EEG signals into 16 levels and the ECoG signals into 16 levels. We then employed joint and

[1] EGI from Electrical Geodesics, Inc. of Eugene, OR; ECoG array from Neuroscan, El Paso, TX

marginal histogram counts to compute the mutual information between each (discretized) EEG-ECoG signal pair g, c as:

$$I(g;c) = \sum_{i=1}^{16}\sum_{j=1}^{16} P(g=i, c=j) \log \frac{P(g=i, c=j)}{P(g=i)P(c=j)} \tag{1}$$

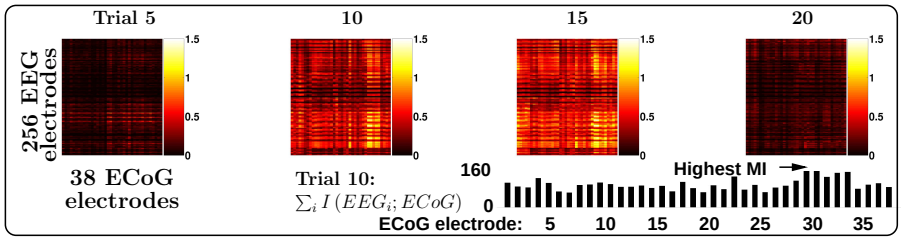

Figure 2: **Mutual information for EEG/ECoG feature selection:** Top plots show average mutual information between each EEG electrode and each ECoG electrode, in number of bits, over 4 different trials of hand movement. At bottom we show the columnwise sum of the top plot for trial 10. This reflects the total amount of information (in bits) about each ECoG signal we expect to recover from the EEG signals. We restrict our regression mapping from ECoG to EEG to consider only the most informative ECoG electrode, number 31 (indicated by an arrow here and in Fig. 1(b)).

Fig. 2 shows pairwise mutual information between all meaningful (i.e. non-ground or reference) ECoG electrodes and all EEG electrodes; some ECoG electrodes are noticeably more informative than others. The bar graph below shows the columnwise sum of one representative matrix, that of trial 10. ECoG electrode 31 appears to be the most informative, and will be used for regression as described below.

3.2 Frequency response

One goal of our analysis is to characterize the frequency response of the ECoG $\rightarrow$ EEG transfer function. This analysis indicates which frequency bands are least attenuated by passing from ECoG to EEG, and may serve to confirm or refute various physiological models that have been proposed [1, 16, 7, 15].

Under the assumption that the matter between cortex and scalp implements a finite impulse response filter, we derived the frequency-domain filter coefficient at frequency f, denoted $\mathcal{F}(f)$, as:

$$\mathcal{A}_\alpha(f) = \sqrt{\tilde{\alpha}(f)\,\tilde{\alpha}^*(f)} \tag{2}$$

$$\mathcal{F}(f) = \frac{\mathcal{A}_{EEG}(f)}{\mathcal{A}_{ECoG}(f)} \tag{3}$$

where $\tilde{\alpha}(f)$ is the Fourier transform of signal α at frequency f, $\tilde{\alpha}^*$ is the complex conjugate of the Fourier transform, and $\mathcal{A}(f)$ is the amplitude (square root of the power) at frequency f. Fig. 3(a) plots frequency response characteristics $\mathcal{F}$ of the channel from ECoG electrode 31 to various EEG electrodes during a single trial. Fig. 3(b) plots mean (black line) and mean plus standard deviation (red dashed line) for the filter coefficients, taken over all EEG electrodes during the single trial.

EEG is susceptible to ambient contamination, and this is reflected by the peaks at 60 Hz. The spectra show heavy attenuation at higher frequencies (greater than $\sim$100 Hz), with the exception of significant power at around 180 Hz. The high intertrial variance (Fig. 3(b)) of the spectrum at that point suggests that the observed power may be attributable to noise in the ECoG recording equipment. Intertrial variance relative to mean signal strength is particularly low for frequencies around 20-50 Hz, suggesting that this frequency band carries high information content.

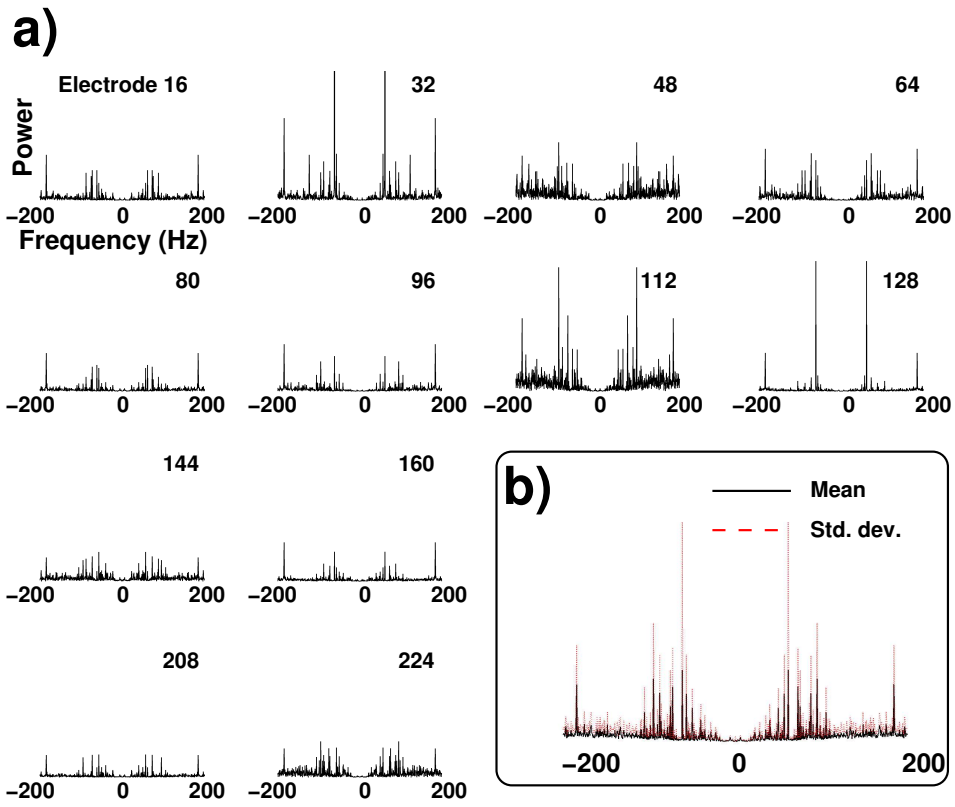

Figure 3: **Channel characteristics from simultaneous EEG/ECoG recording:** (a) Frequency-domain representation of ECoG → EEG transfer function measured between several EEG electrodes and ECoG electrode 31 over the course of a single trial (see text for details). Higher values indicate greater signal power passed from ECoG to EEG; lower values indicate greater attenuation of signals at that frequency band. All plots are scaled identically. (b) Mean (solid black) and standard deviation (red dashed line) of frequency-domain filter coefficients of the ECoG → EEG function, taken over all EEG electrodes. Plot is scaled identically to subplots in (a).

4 Regression techniques

4.1 The forward problem

Many physiological models use basic electrodynamic equations to derive a linear mapping for the forward problem [14, 19, 8]. In this section, we show that an accurate, linear solution to the forward problem can be derived from simultaneous recording without explicit use of electrodynamic equations. Let $\mathbf{C}$ denote the $38 \times t$ matrix of ECoG signals (from time $1 \ldots t$), and $\mathbf{G}$ denote the corresponding $256 \times t$ matrix of EEG signals. We used the pseudoinverse to infer the forward mapping ECoG → EEG:

$$\mathbf{A} = \mathbf{G}\mathbf{C}^{\mathrm{T}} \left(\mathbf{C}\mathbf{C}^{\mathrm{T}}\right)^{-1} \tag{4}$$

The matrix $\mathbf{A}$ is thus a 38×256 matrix that describes how a linear combination of ECoG signals yields each EEG signal. We derived the matrix $\mathbf{A}$ using data from trial 10, and tested using ECoG data from trial 15 to predict EEG responses on trial 15 using the equation $\mathbf{G}_{\mathrm{test}} = \mathbf{A}^{\mathrm{T}}\mathbf{C}_{\mathrm{test}}$. Fig. 4 shows the testing result (true EEG in blue, pseudoinverse prediction in red), demonstrating the viability of linear methods to infer the ECoG → EEG mapping. Note that EEG channels 32 and 128 have high impedance connections with ap-

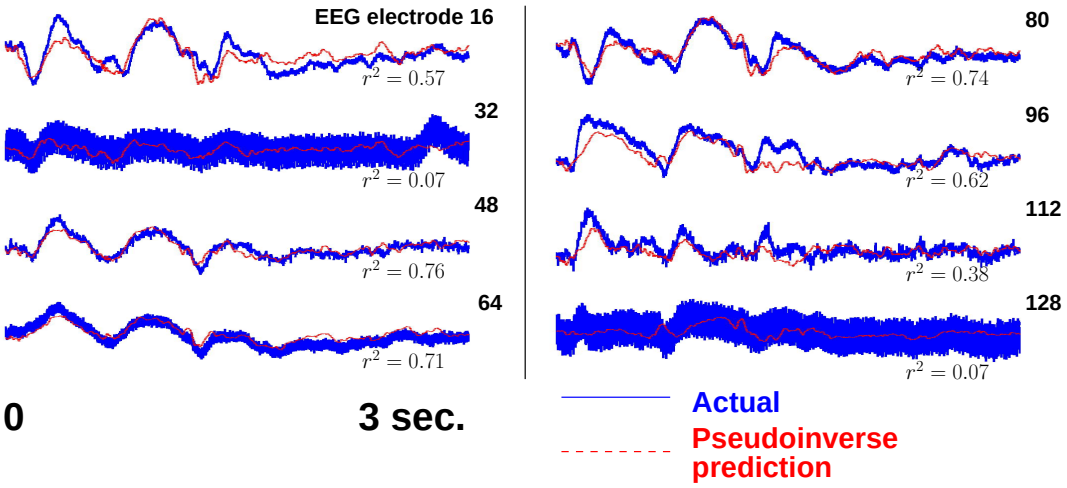

Figure 4: **Linear reconstruction of EEG signals using pseudoinverse:** As expected, simple linear techniques suffice to estimate EEG signals from ECoG. We treat the data as stationary (independent with respect to time), and use the pseudoinverse to compute a matrix $\mathbf{A} \equiv \left(\mathbf{C}^{\mathrm{T}}\mathbf{C}\right)^{-1}\mathbf{C}^{\mathrm{T}}\mathbf{G}$ using the data from a single trial (here, trial 10 of the subject making hand movements). We then test the resulting pseudoinverse mapping using data from another trial (here, trial 15 of the subject making hand movements). The blue solid line is test EEG data for the shown electrodes (on trial 15); the red dashed line is predicted EEG data based on simultaneously recorded ECoG (from trial 15). Below each subgraph we show the correlation coefficients (r^2 values) between the predicted and actual EEG data. Subgraphs are not to scale vertically.

preciable 60 Hz contamination. The pseudoinverse-based forward mapping nonetheless produces a reasonable prediction of the signals for those channels.

4.2 The inverse problem

The inverse problem of mapping EEG $\rightarrow$ ECoG is much more difficult than the forward problem due to temporal and spatial smearing of the cortical signal at the scalp. Here we present some preliminary experiments suggesting the power of statistical machine learning to infer ECoG from EEG. We applied two approaches from statistical machine learning for the inverse mapping: particle filters [18, 9] and Gaussian processes [17]: first, we learned a mapping from EEG to ECoG using a Gaussian process; then we applied the learned GP as an observation model, together with a learned model of ECoG dynamics, to implement a particle filter on ECoG signals. Fig. 5(a) depicts a graphical model corresponding to the GP alone approach (top) and the GP plus particle filter approach (bottom). In the latter, directed arrows between temporally consecutive ECoG readings c show that the particle filter models how the ECoG signal changes over time.

Gaussian processes (GPs) perform a regression mapping from a vector-valued quantity $\mathbf{g} \in \mathbb{R}^n$ (the EEG signal) to a Gaussian-distributed variable c (the ECoG signal) conditioned on $\mathbf{g}$. The conditional probability of c given $\mathbf{g}$ is modeled as a Gaussian: $c \sim \mathcal{N}(\mu_{\mathbf{g}}; \sigma_{\mathbf{g}})$. Here $\mu_{\mathbf{g}}, \sigma_{\mathbf{g}}$ are not the empirical mean and standard deviation of c, but functions of $\mathbf{g}$ determined by a kernel function. A covariance function or *kernel function* k computes the pairwise "similarity" of any two EEG readings $\mathbf{g}, \mathbf{g}'$. We performed Gaussian process regression using two different kernel functions. The first, an isotropic kernel function, uses a single hyperparameter γ for all diagonal covariances; the second uses a vector $\mathbf{a}$ to provide a different covariance for each dimension, and assumes an automatic relevance determination (ARD) distance measure [13].

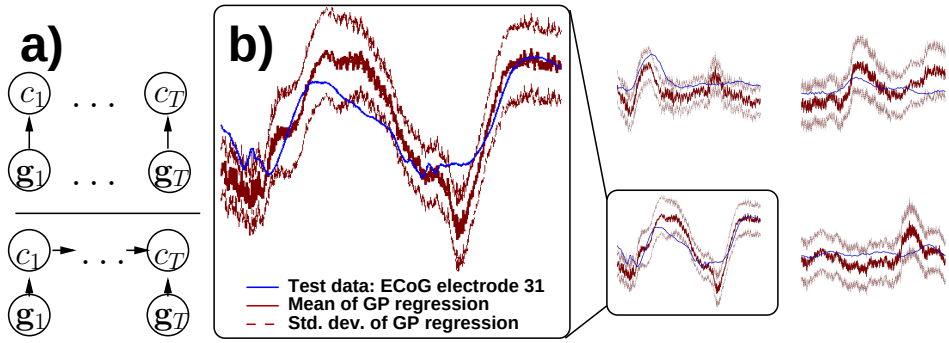

Figure 5: **Gaussian process regression for the inverse problem:** Gaussian processes show some promise in predicting ECoG signals from EEG. (a) Graphical models for the GP alone (top) and GP plus particle filter (bottom). The ECoG signal c generates a vector-valued EEG signal $\mathbf{g}$. The GP plus particle filter model also models the time dependence of ECoG values on previous ECoG values. (b) At left, we show a closeup of 1 second of ECoG testing data from electrode 31 (blue line) plotted against the mean (solid red) and standard deviation (dashed red) of a Gaussian process estimated using an isotropic squared exponential function, and trained using scaled conjugate gradients on the hyperparameters. Training was performed using 1 second of EEG and corresponding ECoG data. At right, we show 1-second testing predictions for several different trials of hand movement (note: vertical axes not to scale).

Particle filters (sometimes called *bootstrap* or *survival of the fittest* filters) approximate arbitrary distributions as collections of point particles, and have been used for diverse applications, including decoding neural signals [11]. On each iteration of the filter, particles are weighted according to a probabilistic observation model (in our case, a Gaussian process with an isotropic squared exponential kernel function). Particles with high weight "survive" to the next distribution, modified by a "proposal" distribution describing the filter dynamics from one iteration to the next. We derived a proposal distribution from ECoG training data by discretizing the ECoG signal into 100 levels, then learning a 100×100 transition kernel describing how the signal changes from time t to time $t+\Delta_t$ (here $\Delta_t = 1$ ms). The particle filter estimate at a given time is given by taking the mean over all particles at that time step. We used the isotropic kernel GP model to weight particles on each time step according to:

$$w_{\hat{\mathbf{c}}} \propto \exp\left(\frac{||\hat{\mathbf{c}} - \mu_{GP}(\mathbf{g})||^2}{\sigma_{GP}(\mathbf{g})^2}\right) \tag{5}$$

where $\hat{\mathbf{c}}$ is a sample from the particle filter over ECoG signals, $\mathbf{g}$ is a vector of EEG electrode measurements, and $\mu_{GP}(\mathbf{g}), \sigma_{GP}(\mathbf{g})$ are the GP-predicted mean and standard deviation of the ECoG signal, conditioned on $\mathbf{g}$.

The GP hyperparameters and particle filter proposal distribution were inferred from 1 second (1000 data points) of data from a single training trial. GP hyperparameters were learned using scaled conjugate gradients on the training ECoG likelihood; the particle filter proposal distribution was computed from transition counts of the discretized ECoG signal. The results shown here are taken from inverse estimates of 1 second of ECoG data on 16 different testing trials.

Table 1 shows mean square error (MSE) reconstructing ECoG from EEG over several testing trials. Here standard deviation represents the standard deviation of the error (residual) values; lower values thus indicate more consistent performance. We normalized across the four methods to show three significant digits. We tested several methods for performing the inverse regression mapping: i) linear (pseudoinverse); ii) Gaussian process regression with

an anisotropic squared exponential kernel and automatic relevance determination (ARD); iii) Gaussian process regression with an isotropic squared exponential kernel; and iv) particle filter with Gaussian process observation model (using an isotropic squared exponential kernel). Note that Gaussian process regression outperforms a linear model by an order of magnitude. Although the GP with anisotropic kernel gives lower MSE, the learned model suffers from two problems. First, learning hyperparameters for the anisotropic kernel failed to converge after 100 iterations of the scaled conjugate gradients algorithm. Second, the prediction from the resulting learned model frequently results in a "flat" waveform, grossly distorting the power spectrum of the resulting signal (see Fig. 6(c) for an example). We suggest that these problems result from the anisotropic kernel overfitting the training data. In contrast, the isotropic kernel seems to preserve spectral characteristics with greater fidelity, with the same order of magnitude MSE as the anisotropic kernel.

Method	**Mean MSE over 16 test trials ($\pm$ mean std. dev.): lower is better**
Linear (pseudoinverse)	96.0 $\pm$ 7.78E-4
GP (isotropic sq. exp. kernel)	7.49 $\pm$ 5.08E-4
GP (anisotropic sq. exp. kernel, ARD)	1.63 $\pm$ 2.32E-4 (convergence problem)
GP (isotropic sq. exp. kernel) + PF	2.40 $\pm$ 3.54E-4

Table 1: Prediction accuracy for the inverse problem: Mean accuracy $\pm$ standard deviation of error residuals for 3 different methods, taken over 16 test trials of 1000 data points each. See text for details.

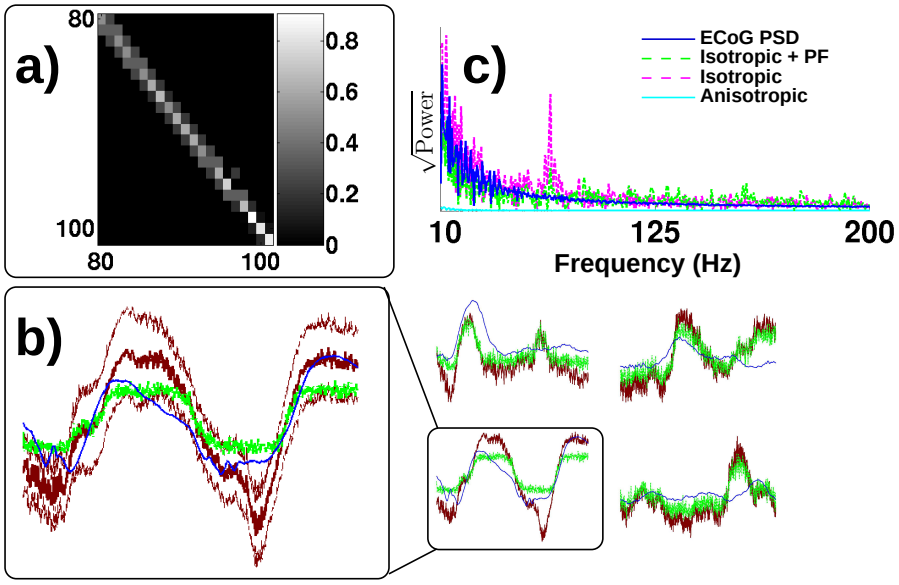

Figure 6: **Particle filter plus GP leads to more accurate inverse estimates:** (a) First-order state transition kernel for ECoG signal. We show the probability of transitioning from some example discretized ECoG values (indexed by rows) to other ECoG values (indexed by columns). High values along the diagonal reflect the slowly-changing nature of recorded ECoG signals; that is, the signal typically stays around the same value it had during the preceding time interval. (b) Examples of ECoG signal prediction from EEG testing data. Shows true ECoG (blue), mean of Gaussian process prediction (red), and particle filter prediction (green). The particle filter estimates are more accurate in terms of both mean square accuracy and (c) the one-sided power spectral density compared to the Gaussian process prediction alone.

Fig. 5(b) shows the ability of a Gaussian process using the isotropic squared exponential kernel to predict ECoG signals from EEG. The solid blue line shows the ground truth ECoG signal; solid red shows the mean of the GP prediction; and dashed red lines show one standard deviation away from the mean. Note that, on most time steps shown in (a), the true ECoG values are within one standard deviation of the GP mean prediction, suggesting that the model has learned to capture at least the gross features of the ECoG waveform.

Fig. 6(a) shows the learned proposal distribution (transition kernel) for the particle filter model. Each row of the matrix is a distribution over next (discretized) states of the ECoG signal at time $t + \Delta_t$, conditioned on the ECoG state at time t. Fig. 6(b) shows in greater detail how a combination of a GP regression model with a particle filter over time can lead to greater accuracy (both in MSE and power spectrum) than GP regression alone. Ground truth ECoG is plotted in solid blue, GP mean prediction in solid red, and the particle filter prediction in dashed green. Fig. 6(c) shows the one-sided power spectrum for the waveform in (b). The GP model captures low-frequency behavior of the ECoG well, but tends to overestimate power at higher frequencies. The combination of GP plus particle filter brings the estimated power spectrum closer to the actual, attenuating power at the higher frequencies while still tracking the lower-frequency components of the waveform with accuracy.

5 Discussion and extensions

Our analysis has focused on information-theoretic criteria for feature selection, and on simple techniques for regression of EEG electrode readings to ECoG readings. Although our results concentrate on time-domain analysis, we anticipate that selecting for features in the frequency domain will also yield valuable models. Future work will concentrate regression methods in the frequency domain. This may be particularly applicable for detecting physiologically relevant signals for brain-computer interface control. Recent simultaneous recordings of EEG and ECoG in an awake, epileptic subject [21] show promising results using a sparser EEG electrode array. The work in [21] employs a finite-element method to infer ECoG from EEG. While their work differs from ours in i) its significantly greater computational complexity; ii) its lack of dense recording arrays; iii) its focus on showing correlation between predicted and actual ECoG (rather than frequency-domain characterization of the predicted and actual ECoG signals); iv) its neglect of ECoG temporal dynamics; and v) its need for extremely accurate head geometry to make predictions, we are intrigued by the possibility of combining their method with ours for more accurate predictions. For example, one might envision a finite-element method where different Gaussian process hyperparameters are learned for each element in a coarse mesh over the head, rather than the single set of hyperparameters learned in the GP here or the fine mesh employed in [21].

Another immediate extension is to generalize our results beyond a single subject. As we have shown, producing a reliable characterization of the EEG-ECoG mapping for a single subject across trials even given ground-truth data is a nontrivial task. For the even more difficult task of generalizing across subjects, we believe Bayesian inference techniques will be of particular value. The attraction of Bayesian inference in the problem of simultaneous recording is the elegant correspondence between distributions and measurements given model parameters θ:

$$P\left(ECoG|EEG\right) = \frac{1}{Z}\int_{\theta} P(ECoG|EEG,\theta)P(EEG,\theta)d\theta \tag{6}$$

Future progress will thus hinge on characterizing the intersubject variance in the $ECoG \rightarrow EEG$ mapping (the higher-order moments of $P(ECoG|EEG,\theta)$), which we expect to have relatively low variance under the assumption of physiological homogeneity in the skull, meninges, etc. A subject-by-subject "burn-in" period while sufficient data is collected to characterize the subject-dependent distribution $P(EEG,\theta)$ will permit infer-

ence of cortical signals. The parameter vector θ could include subject-independent, task-dependent values, such as task context, cortical regions being monitored, etc.

We expect time-series analysis to play an important future role in inferring ECoG data from EEG data; characterizing the evolution of ECoG measurements as a subject performs a particular task remains a challenging future problem. We also expect techniques such as independent component analysis (ICA) [2, 10] to play an important role in deciphering ECoG signals from EEG. Ongoing work in our lab is attempting to correlate independent components of EEG with simultaneously recorded ECoG signals, hoping to conclusively demonstrate the viability of ICA-based approaches to brain-computer interfaces.

Acknowledgements:

We thank Pradeep Shenoy for helpful comments. We benefitted greatly from use of R. Moddemeijer's code to compute mutual information: `http://www.cs.rug.nl/~rudy/matlab/`. Rasmussen's and Williams' code for Gaussian process regression was also instrumental to our efforts: `http://www.gaussianprocess.org/gpml/code/matlab/doc/`. We used the Netlab implementation of scaled conjugate gradients for GP training: `http://www.ncrg.aston.ac.uk/netlab/`.

References

[1] T. I. Alecu, P. Missionnier, S. Volshynovskiy, P. Giannakopoulos, and T. Pun. Soft/hard focalization in the EEG inverse problem. In *IEEE Workshop on Statistical Signal Processing*, pages 17–20, 2005.

[2] A. J. Bell and T. J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1129–1159, 1995.

[3] H. Berger. Über das elektroenkephalogramm des menschen. *Arch. Psychiatr. Nervenkr.*, 1929.

[4] M. S. Berger, J. Kincaid, G. A. Ojemann, and E. Lettich. Brain mapping techniques to maximize resection, safety, and seizure control in children with brain tumors. *Neurosurgery*, 25(5):786–792, 1989.

[5] N. E. Crone, D. L. Miglioretti, B. Gordon, and R. P. Lesser. Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. II. Event-related synchronization in the gamma band. *Brain*, 121(12):2301–2315, 1998.

[6] N. E. Crone, D. L. Miglioretti, B. Gordon, J. M. Sieracki, M. T. Wilson, S. Uematsu, and R. P. Lesser. Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. I. Alpha and beta event-related desynchronization. *Brain*, 121(12):2271–99, 1998.

[7] R. G. de Peralta-Menendez and S. L. Gonzalez-Andino. A critical analysis of linear inverse solutions to the neuroelectromagnetic inverse problem. *IEEE Transactions on Biomedical Engineering*, 45(4), April 1998.

[8] J. Ermer, J. Mosher, S. Baillet, and R. Leah. Rapidly recomputable EEG forward models for realistic head shapes. *Phys Med Biol.*, 46(4):1265–1281, Apr 2001.

[9] N. J. Gordon, D. J. Salmond, and A. F. M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proceedings on Radar and Signal Processing*, 140:107–113, 1993.

[10] C. Jutten and J. Herault. Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24:1–10, 1991.

[11] R. Kelly and T. S. Lee. Decoding V1 neuronal activity using particle filtering with Volterra kernels. In *Advances in NIPS*, 2003.

[12] E. C. Leuthardt, G. Schalk, J. R. Wolpaw, J. G. Ojemann, and D. W. Moran. A brain-computer interface using electrocorticographic signals in humans. *J. Neural Eng.*, 1(2):63–71, 2004.

[13] D. J. Mackay. Bayesian interpolation. *Neural Computation*, 4:415–447, 1992.

[14] G. Marin, C. Guerin, S. Baillet, L. Garnero, and G. Meunier. Influence of skull anisotropy for the forward and inverse problem in EEG: simulation studies using FEM on realistic head models. *Human Brain Mapping*, 6(4), 1998.

[15] J. C. Mosher and R. M. Leahy. Recursive MUSIC: a framework for EEG and MEG source localization. *IEEE Transactions on Biomedical Engineering*, 45(11), November 1998.

[16] R. D. Pascual-Marqui, M. Esslen, K. Kochi, and D. Lehmann. Functional imaging with low resolution brain electromagnetic tomography (LORETA): a review. *Methods and Findings in Exp. and Clin. Pharmacology*, 24C:91–95, 2002.

[17] C. E. Rasmussen and C. K. I. Williams. *Gaussian processes for machine learning*. MIT Press, 2006.

[18] D. B. Rubin. Using the SIR algorithm to simulate posterior distributions. In M. Bernardo, K. DeGroot, D. Lindley, and A. Smith, editors, *Bayesian Statistics 3*. Oxford, UK: Oxford University Press, 1988.

[19] N. von Ellenrieder, C. Muravchik, and A. Nehorai. A meshless method for solving the EEG forward problem. *IEEE Trans Biomed Eng.*, 52(2):249–257, Feb 2005.

[20] J. R. Wolpaw, D. J. McFarland, G. Neat, and C. Forneris. An EEG-based brain-computer interface for cursor control. *Electroencephalogr. Clin. Neurophysiol.*, 78(3):252–259, 1991.

[21] Y. Zhang, L. Ding, W. van Drongelen, K. Hecox, D. M. Frim, and B. He. A cortical potential imaging study from simultaneous extra- and intracranial electrical recordings by means of the finite element method. *NeuroImage*, April 2006.