Analysis of Neuronal Source Dynamics During Seizure Using Vector Autoregressive Models, ICA, Sparse Bayesian Learning and ECoG

Tim Mullen
Swartz Center for Computational Neuroscience
Institute for Neural Computation and
Department of Cognitive Science
UC San Diego
La Jolla, CA
tim@sccn.ucsd.edu

Abstract

Accurate detection of seizure onset as well as identification of neuronal regions critically involved in initiating and propagating a seizure remains an important area of research. Understanding the dynamics of neural processes underlying different stages of a seizure can help in devising novel methods of seizure detection, intervention and treatment. In this paper we analyze linear neuronal dynamics during epileptic seizures using adaptive multivariate autoregressive (VAR) models applied to maximally-independent (ICA) sources of intracranial EEG (iEEG, ECoG) data recorded from subdural electrodes implanted in a human patient for presurgery monitoring. We analyze the time-frequency dynamics of directed information flow between sources using a multivariate granger-causal method (dDTF), identifying distinct patterns of information flow in different stages of the seizure. We then further examine the spatial distribution in the cortical source domain of causal sources and sinks of ictal activity using a novel combination of causal flow metrics and Sparse Bayesian Learning-based source localization. Finally, we apply an eigendecomposition method to decompose the VAR model into a system of decoupled oscillators and relaxators (eigenmodes) with characteristic damping times and frequencies. We demonstrate that analysis of a small subset of the most dynamically important eigenmodes may allow effective detection of ictal onset and offset, while also yeilding insight into the dynamical structure of the neuronal system.

1 Introduction

Epilepsy is one of the most common neurological disorders, affecting 50 million people worldwide, and in approximately 30% of these patients the seizures are not controlled by any available medical therapy. About 4.5% of all patients with epilepsy are thus potential candidates for surgical treatment. In this patient group, surgery may have a good chance of success, but only if the brain region(s) generating seizures can be accurately localized. For this purpose, in selected cases, recordings are acquired using subdural and/or depth electrode (intracranial) pre-surgical evaluation. The aim of this paper is to describe some recent preliminary assays in modeling and analyzing the spatial and time-frequency dynamics of seizure generation and propagation from intracranial EEG recordings using data acquired by Dr. Worrell at the Mayo Clinic. Frequency-domain measures have proven useful for studying seizure dynamics due to the inherent oscillatory structure present in ictal activity. Under suitable conditions, a vector autoregressive (VAR) model provides an idealized model for the

analysis of oscillatory structure in stochastic time series [7, 6, 17, 24]. From the VAR coefficients, we can obtain a number of useful quantities describing the power spectral modulation and information flow between neuronal process. One such measure, Granger-causality, has received much attention in recent years due to its relative simplicity and proven usefulness in identifying directed oscillatory information flow in neuronal systems (see [4] for a review). Several previous applications of VAR modeling to identifying frequency-specific Granger-causal influences during seizure have used the normalized Directed Transfer Function (DTF) measure (e.g. [10, 22]). This measure suffers from several problems, including the inability to distinguish between direct and indirect causal influences (thus it does not truly describe Granger-causality in a multivariate sense). Furthermore, the causal influences are normalized by the inflows to the causal source, making identification of causal sinks difficult, as well as rendering potentially misleading results when directly comparing DTF values between different pairs, or across different times or frequencies. Here we use a modified version of the DTF, the short-time direct DTF (SdDTF) which measures only direct influences (within the system of observed variables) and furthermore normalizes over all outflow and inflow, across all time and frequency, making it possible (assuming equal variance in the signals) to directly compare the amplitude of the causal measure between different pairs of variables at different time points and frequencies.

In order to identify anatomical regions critically participating in seizure generation and propagation, it is important to be able to localize the sources of observed EEG or iEEG. In this paper we will extend a previous report of the use of a powerful new distributed source localization algorithm, based on Sparse Bayesian Learning (SBL), to localize maximally-independent sources of iEEG activity [1, 2]. We will demonstrate the use of these inverse solutions to visualize the time-evolving distribution of transient information outflow and inflow on the cortical surface for one or more IC sources.

Finally, we will apply an eigendecomposition method to decompose the VAR model into a system of decoupled oscillators and relaxators (eigenmodes) with characteristic damping times and frequencies. Our objective here being to demonstrate that analysis of a low-dimensional subset of the most dynamically important eigenmode may allow effective detection of ictal onset and offset, while also yeilding insight into the dynamical structure of the neuronal system.

2 Theory

2.1 Independent Component Analysis

Infomax ICA is a method for blind separation of an M-dimensional process into N maximally independent sources under the assumption of nongaussianity of the sources [3]. This method has been shown to be effective at identifying and separating functionally-independent neuronal processes while increasing signal-to-noise through removal of eye, muscle, and electrical or channel noise artifacts, as well as volume-conducted potential mixtures from the data [15]. The basic framework for ICA, as applied to EEG source separation is as follows:

Suppose we have an $M \times T$ matrix $X = x_1, \ldots, x_T$ consisting of M channels and T time points. Let us assume that the observed signals at time t, $x_t = [x_t^{(1)}, \ldots, x_t^{(M)}]^T$, are linear, instantaneous mixtures of $N \leq M$ unobserved nongaussian sources, $s_t = [s_t^{(1)}, \ldots, s_t^{(N)}]^T$. The generative model for an observed signal vector x_t can thus be written as:

$$x_t = As_t \tag{1}$$

where A is an unknown $M \times N$ nonsingular "mixing matrix." Our goal is to identify both s_t and A, given only x_t, with the objective of maximizing the statistical independence of the sources $S = s_1, \ldots, s_T$. More formally, we wish to identify an "unmixing" matrix $W = A^{-1}$ such that.

$$\tilde{s}_t = Wx_t \tag{2}$$

where $\tilde{s}_t \equiv s_t$ up to some scaling and/or permutation and where the joint probability density function of the sources factorizes: $P(s^{(1)}, \ldots, s^{(N)}) = \prod_{k=1}^{N} P(s^{(k)})$. There are a number of solutions to the

problem; however, a pratically robust solution is given by the extended Infomax approach of Bell, Sejnowski and Lee [3, 12]. Here W is obtained by a stochastic gradient ascent approach with the update equation:

$$\begin{aligned} \Delta W &= [I - K\tanh(\tilde{S})\tilde{S}^T - \tilde{S}\tilde{S}^T]W \\ K_{ij} &= \delta_{ij} f(\tilde{s}_i) \end{aligned} \tag{3}$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. $f(\tilde{s}_i)$ is -1 if $\tilde{s}_i$ is subgaussian and 1 if $\tilde{s}_i$ is supergaussian. K can be derived from the generic stability analysis of the separating solutions. Having obtained the solution, we will denote the columns of W *Independent Components (ICs)* and the time course of the estimated sources $\tilde{s_t}$ *IC activations*.

2.2 Adaptive multivariate autoregressive modeling

Assuming that $X = [x_1, \ldots, x_T]$ is an M-dimensional zero-mean weakly-stationary stochastic process of length T, we can describe the linear dynamics of the state vector $x_t = [x_t^{(1)}, \ldots, x_t^{(M)}]^T$ as a vector autoregressive (VAR[p]) process of (possibly infinite) order p:

$$x_t = \sum_{l=1}^{p} A_l x_{t-l} + u_t \tag{4}$$

where $u_t \in \Re^{M \times 1}$ is a zero-mean white noise process with covariance matrix

$$\Sigma = \left\langle u_t u_t^T \right\rangle$$

The coefficient matrices, A_l, can be estimated using a number of approaches, including multivariate ordinary and stepwise least-squares approaches, lattice algorithms (e.g. Vieira-Morf) or state-space models (Kalman filtering) [19, 14]. Neumaier and Schneider [19] provide an efficient stepwise least-squares algorithm which we use in this paper.

For non-stationary data, we can model the time-varying dynamics using a simple segmentation approach [11, 9]. We fit separate VAR[p] models to a sequence of overlapping locally-stationary windows of length W. We generally choose a small step size, $Q \ll W$, yeilding highly-overlapping windows so that coefficient matrices vary smoothly with time. This approach yields $\lfloor \frac{T-W}{Q} + 1 \rfloor$ VAR[p] models, each of which describes the local linear dynamics of the process, within the respective window.

2.2.1 Spectral measures and Granger-causality

Electrophysiological processes generally exhibit oscillatory structure, making them well suited for frequency-domain analysis [8]. A suitably fit autoregressive model provides an idealized model for the analysis of oscillatory structure in stochastic time series [7, 6, 17, 24]. From the AR coefficients, we can obtain a number of useful quantities including the spectral density matrix and the transfer function of the process. From these and related quantities we can obtain power spectra, coherence and partial coherence, Granger-Geweke causality, directed transfer function, and a number of other quantities increasingly being used by the neuroscience community to study synchronization and information flow in the brain.

To obtain our frequency-domain representation, let us begin with our VAR[p] model (Eq. 4). Rearranging terms:

$$u_t = \sum_{l=0}^{p} \hat{A}_l x_{t-l} \;\text{ where }\; \hat{A}_l = -A_l \;\text{ and }\; \hat{A}_0 = I. \tag{5}$$

Z-transformation yields:

$$\begin{aligned} U(f) &= A(f)X(f) \\ A(f) &= \sum_{l=0}^{p} \hat{A}_l e^{-i2\pi f l} \end{aligned} \tag{6}$$

The $(M \times M)$ spectral density matrix of the process is given by

$$S(f) = X(f)X(f)^* = H(f)\Sigma H(f)^{-1} \tag{7}$$

where * denotes conjugate transpose and

$$X(f) = A(f)^{-1}U(f) = H(f)U(f). \tag{8}$$

The (squared) Short-time Direct Directed Transfer Function (SdDTF) [**?**] from process j to i is given by

$$\eta_{ij}^2(f,t) = \frac{|H_{ij}(f,t)|^2 |P_{ij}(f,t)|^2}{\sum_{f,\tau}\sum_{kl}|H_{kl}(f,\tau)|^2|P_{kl}(f,\tau)|^2} \tag{9}$$

where $P_{ij}(f) = \dfrac{\hat{S}_{ij}(f)}{\sqrt{\hat{S}_{ii}(f)\hat{S}_{jj}(f)}}$, $\hat{S} = S^{-1}$ is the partial coherence between variables i and j.

The SdDTF for time window centered at time t, $\eta_{ij}^2(f,t)$, can be interpreted as a normalized measure of the direct granger-causal influence at time t from process j to i at frequency f, conditioned on all other measured variables.

2.2.2 Graph-theoretic measures

For a given time-frequency tuple, (t, f), we can represent the causal dynamics of the multivariate neural system as a directed graph where the M neural sources correspond to nodes and a significant value of η_{ij}^2 corresponds to a directed edge from node j to node i. The causal participation of node j within the rest of the system can be represented by the *outflow*, *inflow*, and *causal flow*:

$$\textit{Outflow: } \Omega_j = \sum_{i=1}^{M}\eta_{ij}^2, \quad \textit{Inflow: } \Upsilon_i = \sum_{j=1}^{M}\eta_{ij}^2, \quad \textit{Causal flow: } F_i = \Omega_i - \Upsilon_i \tag{10}$$

Outflow characterizes the causal influence of a node on the rest of the system, while the degree to which a node is causally driven by other elements of the system is represented by the inflow. The causal flow represents the asymmetry in causal influence of a given node. Large positive values of F_i indicate a causal source (hub) while large negative values indicate a causal sink. Values near zero indicate balanced inflow and outflow or nonsignificant flow.

2.2.3 Decomposition of a dynamical system into eigenmodes

Using the eigendecomposition method of Neumaier and Schneider [17], it can be shown that a stable M-dimensional VAR[p] model can be decomposed into Mp, M-dimensional decoupled eigenmodes, which can each be characterized as an oscillator or relaxator with a characteristic frequency and damping time. The dynamics of the eigenmodes can be described by a system of Mp univariate VAR[1] models coupled only by the covariance of the noise terms. Analysis of the eigenmodes can provide insight into the linear dynamics of the system under observation.

In brief, we begin by noting that the VAR[p] process described in equation 4 is equivalent to the VAR[1] process.

$$\tilde{x}_t = \tilde{A}\tilde{x}_{t-1} + \tilde{u}_t \tag{11}$$

with augmented state vectors and noise vectors

$$\tilde{x}_t = \begin{pmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-p+1} \end{pmatrix} \in \Re^{Mp} \quad \text{and} \quad \tilde{u}_t = \begin{pmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \Re^{Mp}$$

and with coefficient matrix

$$\tilde{A} = \begin{pmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & \cdots & I & 0 \end{pmatrix} \in \Re^{Mp \times Mp} \tag{12}$$

and singular noise covariance matrix

$$\tilde{\Sigma} = \langle \tilde{u}_t \tilde{u}_t^T \rangle = \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \in \Re^{Mp \times Mp}$$

Note that $\tilde{x}_t$ represents a delay embedding of the original state vectors x_t. If $\tilde{A}$ is nonsingular then $\tilde{A} = Q\Lambda Q^{-1}$ where the columns of Q are the eigenvectors (eigenmodes) of $\tilde{A}$ and $\Lambda = \text{diag}(\lambda_k),\ k = (1, \ldots, Mp)$ is the associated diagonal matrix of eigenvalues. The original state and noise vectors can then be represented as

$$\tilde{x}_t = Q\tilde{x}'_t,\ \ \tilde{u}_t = Q\tilde{u}'_t \tag{13}$$

with eigenmode coefficient vector $\tilde{x}'_t = [\tilde{x}_t^{(1)}, \ldots, \tilde{x}_t^{(Mp)}]^T$ and noise vector $\tilde{u}'_t = [\tilde{u}_t^{(1)}, \ldots, \tilde{u}_t^{(Mp)}]^T$. Note that $\tilde{x}'_t = Q^{-1}\tilde{x}_t$ is a rotation of the delay-embedded state vectors into the coordinate system of the eigenvector basis. Substituting these expansions into Eq. 11 for the VAR[1] model, and using the diagonality of Λ, we can represent the coefficient vectors, $\tilde{x}'_t$ as a system of univariate VAR[1] models:

$$\tilde{x}_t^{(k)} = \lambda_k \tilde{x}_{t-1}^{(k)} + \tilde{u}_t^{(k)},\quad k = \{1, \ldots, Mp\} \tag{14}$$

which are coupled only via the transformed, augmented covariance matrix of the noise coefficients: $\tilde{\Sigma}' = Q^{-1}\tilde{\Sigma}Q^{-*}$. In the complex plane, the expected values of the eigenmode coefficients describe a spiral $\langle v_{t+l}^{(k)} \rangle = \lambda_k^l \langle v_t^{(k)} \rangle = e^{-k/\tau_k} e^{(arg\lambda_k)il} \langle v_t^{(k)} \rangle$

with *damping time*

$$\tau_k = \frac{-1}{F_s \ln |\lambda_k|} \tag{15}$$

and *characteristic frequency*

$$f_k = \frac{F_s |\arg \lambda_k|}{2\pi} \tag{16}$$

Here F_s denotes the samping rate of the time series. The damping time (also known as the *e-folding* time [20]) denotes the time required (here in units of seconds) for an initial amplitude $|\tilde{x}_0^{(k)}| = q$ to decay to $|\tilde{x}_\tau^{(k)}| = q/e$. As Von Storch notes in his review on POP analysis [20], an eigenmode analysis using the linear, stationary model (Eq. 11) preferentially "sees" an oscillation in its mature state when noise is relatively small and damping is due to nonlinear and other, unobserved, processes. The damping time provides a statistical measure of how long, on average, the signal is seen before stochastic noise, as well as unobserved or nonlinear dynamical processes become more and more important. A given eigenmode can be characterized as a stochastically forced *oscillator* or *relaxator* based on the sign and/or reality of the associated eigenvalue. Figure 1 shows the classification, along with the characteristic frequency, of an eigenmode with an associated positive, negative, real, or complex eigenvalue. For a stable VAR model with nonsingular coefficient matrix $\tilde{A}$, the modulus of all eigenvalues lie between 0 and 1 and thus the damping time for all eigenmodes is positive and bounded. The variance of the amplitudes of the k^{th} eigenmode coefficients (*excitations*) $\sigma_k = \langle |\tilde{x}_t^{(k)}|^2 \rangle = \tilde{\Sigma}'_{kk}/(1 - |\lambda_k|^2)$ can be interpreted as the dynamical importance of the k^{th} eigenmode $Q_{:k}$. Analysis of the most dynamically important eigenmodes can help elucidate the global dynamical structure of the system.

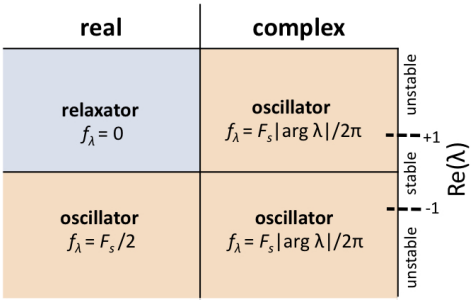

Figure 1: Characterization of eigenmodes of a VAR[p] process. The sign and reality of an eigenvalue λ determines whether the associated eigenmode is an oscillator or relaxator. f_λ is the characteristic frequency of the mode and F_s is the process sampling rate.

3 VAR modeling of seizure dynamics

Intracranial EEG was collected from a patient undergoing presurgical evaluation at the Mayo Clinic (Rochester, MN). The patient presented with seizures due to a porencephalic cyst in the fronto-parietal brain. Sixteen minutes of 78-channel iEEG data (Figure 2) was collected at a sampling rate of 500 Hz during drowsy resting. The data contained two seizure bursts, each lasting around 2 minutes. The 78-channel data was decomposed by extended Infomax ICA into 78 maximally independent component (IC) processes. Through visual inspection, 16 ICs were identified as exhibiting clear epileptiform activity and the remaining ICs were discarded. The selected "seizure" ICs were then localized using an anatomically-realistic head model and a Sparse Bayesian Learning (SBL) algorithm. In brief, a realistic individual head model was constructed for this patient using structural models of graymatter, whitematter, CSF, skull, and scalp, extracted from pre-surgical MR and post- surgical CT images. The forward problem of electromagnetic source localization was solved using the Boundary Element Method (BEM). Source localization was then performed for each IC by applying an SBL method [23], and using the respective column of the inverse of the ICA weight matrix (IC "grid map") as the observation vector. For more details see [2, 1].

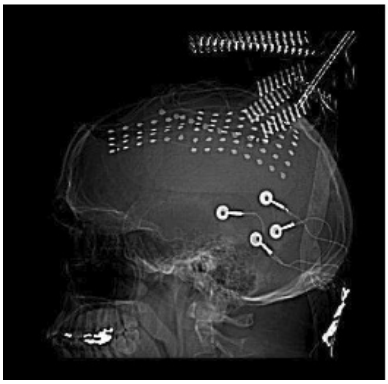

Figure 2: CT image of the implanted grid electrodes. The two grids (6 8, 46) and one medial strip (18) implanted in the patient for monitoring.

The time course of the ICs (activations) were downsampled to 256 Hz using a zero-phase FIR antialiasing filter. Each IC activation sequence was then independently normalized by subtracting the temporal mean and dividing by the temporal standard deviation. This ensures each time series has zero mean and variance one. A 16-dimensional VAR[7] model was fit to the normalized IC activations using the ARFIT stepwise least-squares algorithm [19]. An adaptive model was realized using a sliding window of length 15 seconds, with a step size of 1 second. The model order (p=7) was selected based on inspection of the distribution, over all windows, of model orders that minimized the Hannan-Quinn and SBC information criteria. For each window, the spectral density, coherence, and SdDTF estimators were obtained from the model coefficient and noise covariance matrices, as

described in Section 2.2.1. The outflow, inflow, and causal flow (Eq. 10) were also computed for each IC source. Finally, the VAR[7] model was subjected to an eigendecomposition, from which was obtained the damping times and characteristic frequencies of the 112 eigenmodes.

While it may seem physical interactions (e.g., Granger-causality) between and statistical independence of two systems (as assumed by ICA) are contradictory assumptions, it can be shown that for weakly coupled (e.g., partially coherent or transiently coupled) systems they may be reconciled, since the amplitude distribution of weakly interacting sources may still be statistically independent (or near-independent) [18].

3.1 Results

Figure 3 shows the time course of activations of the selected ICs during onset (left) and offset (right) of the first seizure. Note that IC12 appears to demonstrate earliest onset of ictal activity, followed closely by ICs 13, 11 and 1. The seizure terminates abruptly at 349.5 seconds.

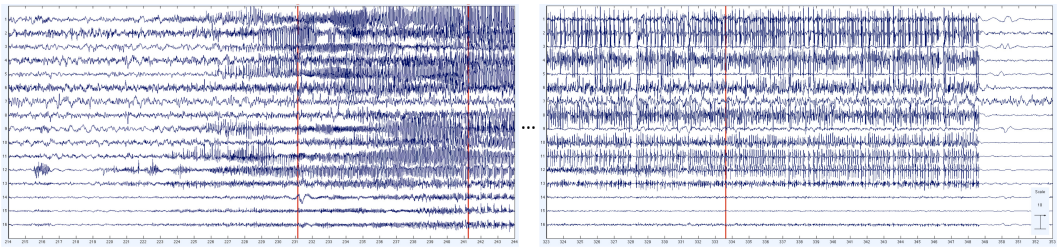

Figure 3: Time course of activations of selected ICs during seizure 1 onset (left) and offset (right). Time units are in seconds.

3.1.1 Stability analysis

A VAR[p] model is stable if the roots of its reverse characteristic polynomial lie outside the unit circle. This is equivalent to all eigenvalues of $\tilde{A}$ having modulus less than 1 [14]. Figure 4 plots the stability index $\varsigma = \ln|\max_k \lambda_k|$ of the fitted model for each window. Note that the process is stable for all time, but starts to loose stability in the first part of the seizure becomes highly stable in the mid-end of the seizure and again approaches instability in the inter-ictal and post-ictal periods. The dramatic decrease in stability following ictal onset may be due to nonlinear dynamics dominating as the system bifurcates from one stable attractor to another. The increase in stability towards the middle of the seizure may reflect a shift towards a high-amplitude, highly-sychronized state of the system exhibiting stable limit cycles. The low stability in the inter-ictal period may reflect the propensity for the system to again bifurcate into a seizure state.

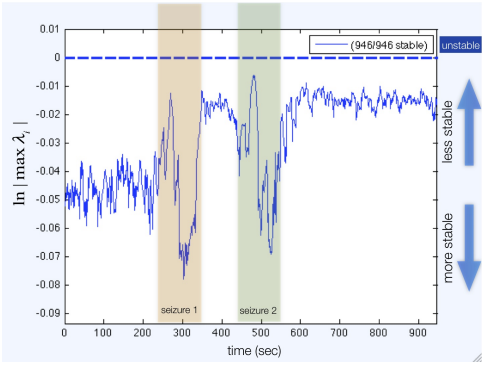

Figure 4: Results of stability analysis. Colored regions denote ictal periods

3.1.2 Causality analysis

Figure 5 shows a *Time Frequency Grid* indicating the spectrum (diagonal) and SdDTF (off-diagonal) interactions between all IC sources for a range of window centers (7.5 to 575 sec) and frequencies (1 to 70 Hz with 1 Hz resolution). Here we have shown the deviation of each measure, for each frequency, from the average value at that frequency within a pre-ictal baseline window of 0 to 107.5 seconds (redshift = above baseline mean, blueshift = below baseline mean). The distribution of each source on the cortical surface, as estimated by SBL, is shown on the column marginal. Information flows from the source indicated on the column marginal to that indicated on the row marginal. Note the prominent bursts of theta (5 Hz), alpha (12 Hz), and beta (12-20 Hz) information flow occuring during different stages of each seizure (each seizure is marked by red (start) and black (end) vertical lines). ICs 1, 5, and 9 appear to be prominent outflow hubs exerting strong influence on multiple other elements of the network in different stages of the seizure. IC1 and IC9 show bursts of outflow primarily at ictal onset and offset while IC5 becomes a causal source primarily in the mid-seizure and at ictal offset.

Interestingly, IC12 was identified as the epileptogenic focus due to it exhibiting prominent inter-ictal discharges (IEDs) and exhibiting ictal onset prior to all other sources. However, this region does not appear to be a strong causal outflow hub, but rather it participates in bidirectional causal interaction with IC1 early in the seizure and (only weakly) with IC9, IC3, and IC5 in mid-seizure. IC12 does however appear to be a causal inflow hub in the mid-late and post-ictal periods. This suggests that regions exhibiting early ictal activity may not be the most prominent sources of seizure propagation. Thus identifying epileptogenic focii based solely on the time-course of ictal activity (including examining power spectra) may produce an incomplete, or even misleading, picture.

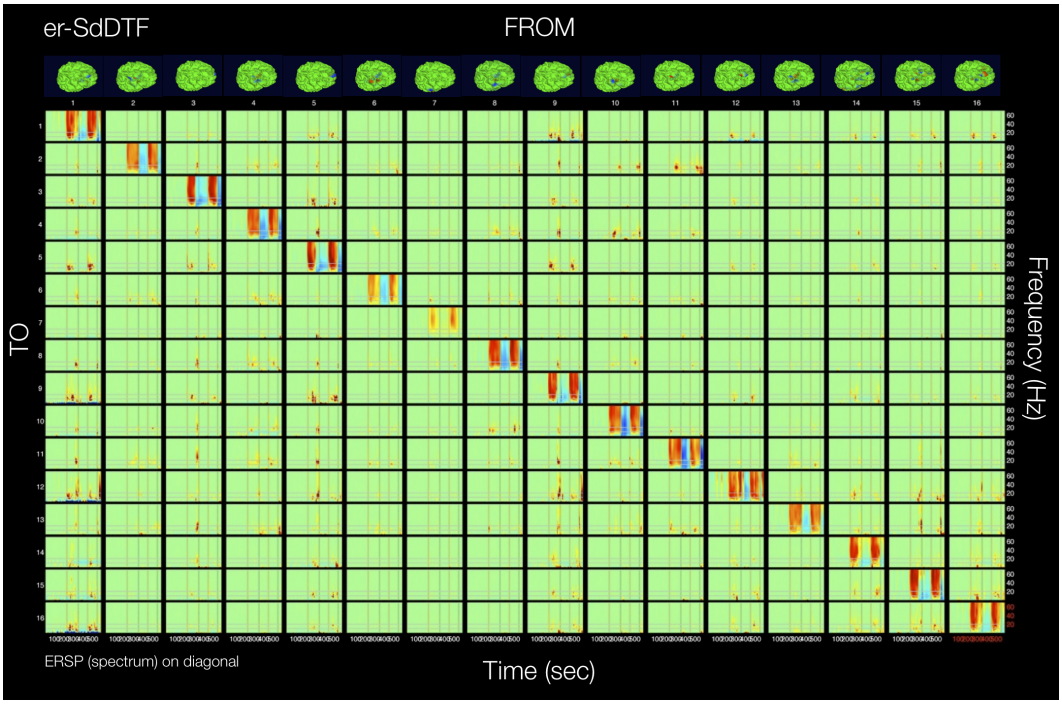

Figure 5: Time-Frequency Grid showing event-related SdDTF (baseline = [0 100] sec) information flow between IC sources. Information flow from columns to rows. Event-related spectral perturbation (ERSP) of each source is shown on the diagonal. Vertical red (black) dashed lines indicate start (end) of seizures. Horizontal gray lines denote 5, 10, 20 Hz markers.

A seizure may be propagated through sychronized activity within and between local networks, connected by long-range cortico-cortical (or subcortical) connections. Here we sought to examine the interactions, at different stages of the seizure, between several IC sources localized to prefrontal and posterior frontal/anterior parietal (precentral gyrus) cortex. Fig. 6 shows a montage containing several of these sources. Here we see three distinct stages of the seizure involving two spatially-distal

sub-networks, presumably connected via cortico-cortical or subcortical white matter tracts. In the first stage of the seizure, we see strong influences from two coupled sources in the prefrontal network (ICs1,9) to a third prefrontal source (IC5). In the second stage of the seizure we see a reversal of the flow, with feedback from IC5 to ICs 1 and 9. In the third stage of the seizure we see all elements of the prefrontal network strongly influencing sources in the precentral network. Finally, following the seizure we see some continued synchronization within the precentral network before the system settles into a damped state. The second seizure shows a similar pattern of propagation, but with reduced flow into the precentral network in the third stage.

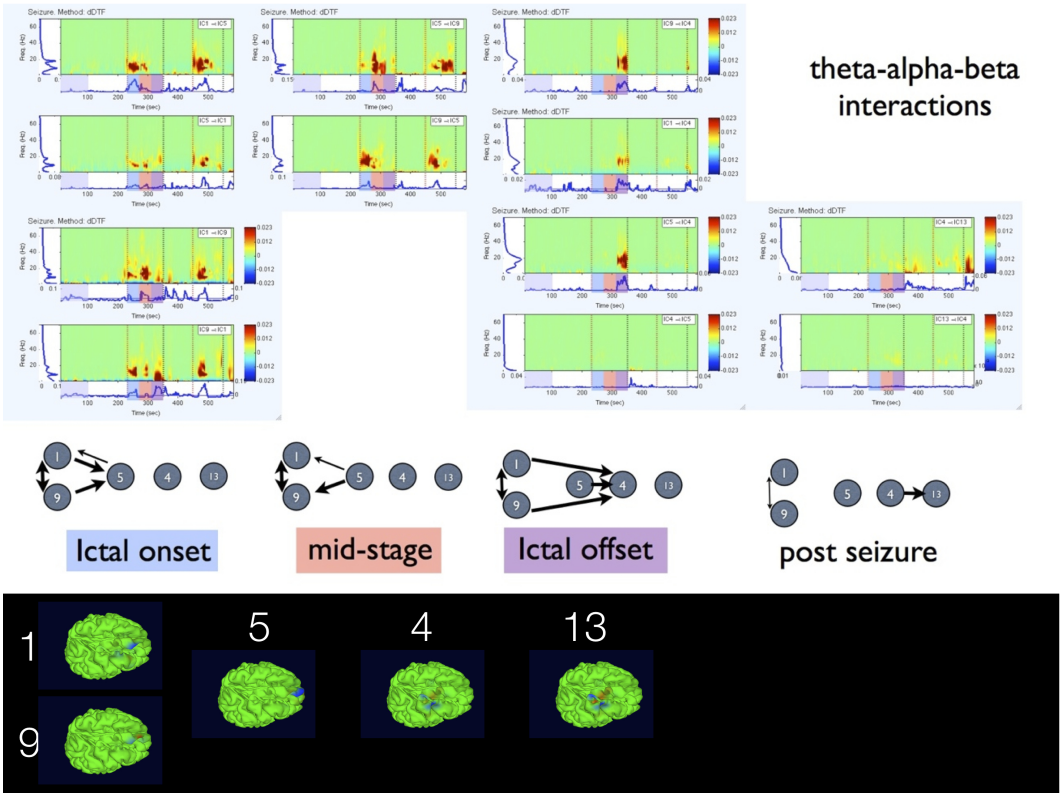

Figure 6: Montage showing dDTF interactions for five selected source during different stage of the seizure(s).

In order to better visualize the time course and spatial distribution of subnetworks critically involved in seizure propagation (causal hubs), we integrated information flow from 2 to 30 Hz and projected the outflow and causal flow for each IC into the source domain using the inverse solution obtained by SBL. In brief, where one would normally multiply channel potentials or component activations by the SBL-learned inverse solution to obtain the time course of potentials for each dipole, here we replace the time course of activations with the time course of Ω_i or F_i (Eq. 10) and multiply it by the absolute value of the inverse solution to obtain the distributions of outflow or causal flow for a specified IC source across the patch basis learned by SBL. Thus, a large projected outflow value at a given voxel can be taken to reflect a large amount of outflow from one or more distributed sources underlying that voxel.

A sequence of frames from two movies, showing outflow and causal flow during the first seizure, are shown in Figure 7. Note the alterating pattern of feedforward (mid-frontal (IC1,9) $\rightarrow$ frontopolar (IC5)) and feedback activity in the first two stages of the seizure. The final stage of the seizure is characterized by strong prefrontal $\rightarrow$ precentral activity spreading from frontopolor (e.g. IC5) and later mid-frontal cortex, posteriorly to posterior frontal and anterior parietal cortex. While the causal flow measure more clearly delineates the causal hubs throughout the seizure, it is important to remember that it cannot distinguish between zero flow (decoupled network) and balanced flow

(symmetric information flow) and therefore should always be examined in combination with another measure such as outflow.

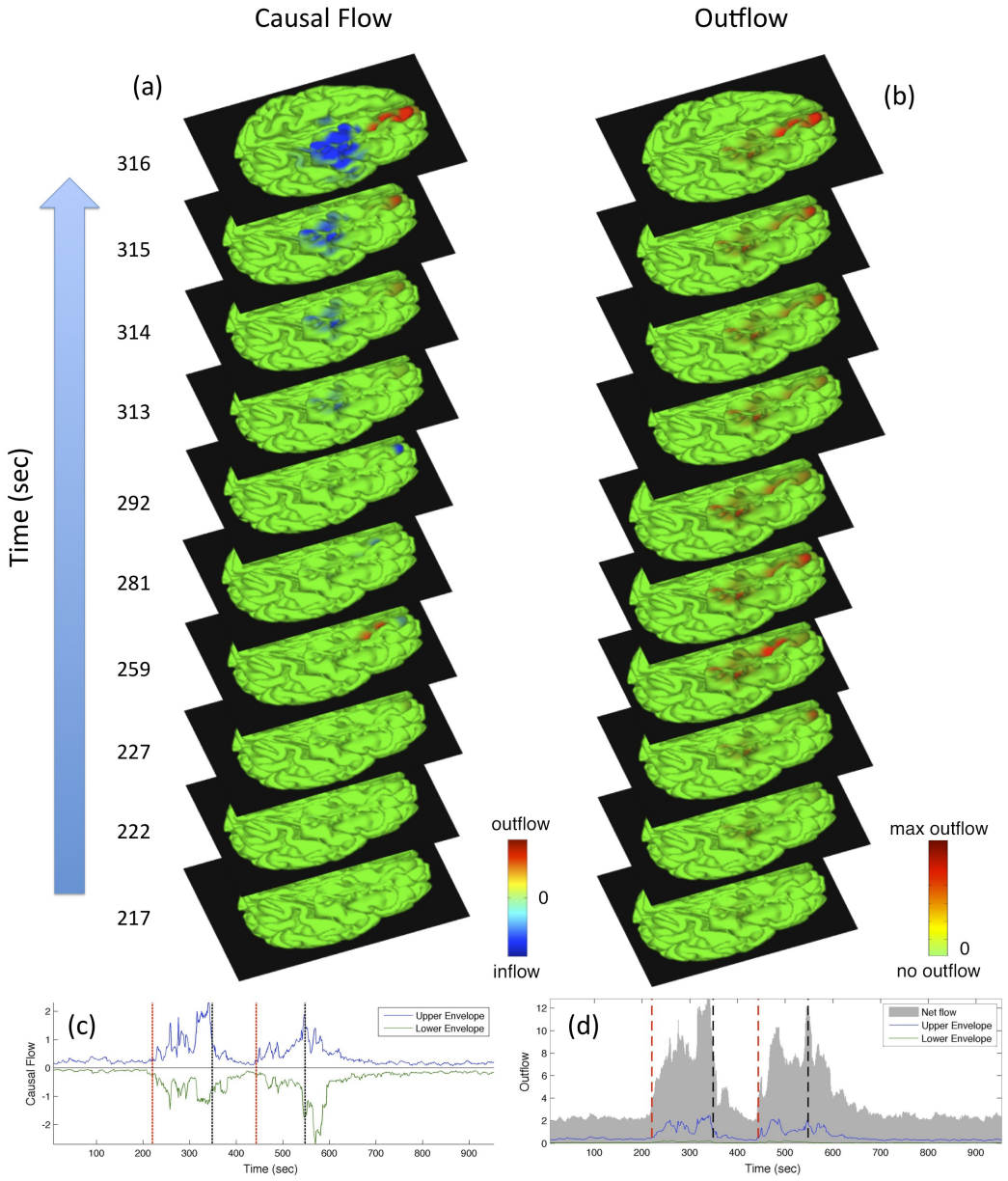

Figure 7: A sequence of frames from a movie showing Causal Flow (a) and Outflow (b) in the source domain at different stages of the seizure. Panel (c) shows the envelope of the causal flow as a function of time. Panel (d) shows the envelope of the outflow, as well as the net outflow (summed outflow over all nodes) as a function of time. Red (Black) dashed lines denote seizure onset (offset)

3.1.3 Eigenmode analysis

Identifying causal hubs may provide a useful way to isolate cortical areas critical to seizure propagation or termination. However, causal hubs are also ubiquitous in normal brain function [5]. Additional techniques for identifying the onset of a seizure may be combined with anatomically-localized causal analysis to help reduce the chance of false-positives. The onset of a seizure is characterized by a shift in the system dynamics from a relatively decoupled state to a highly synchronized state

wherein most elements of the system exhibit a stable limit cycle (oscillation) with some characteristic frequency. Identifying a transition of the global dynamics of the system into such an oscillatory state may be a useful mechanism for seizure prediction. As we discussed in section 2.2.3, a stable M-variate VAR[p] process can be decomposed into Mp decoupled oscillators and relaxators (eigenmodes) each possessing a characteristic frequency (which is zero for relaxators) and damping time, as well as a measure of the eigenmode's dynamical importance (*excitation*). These elements are coupled only by the covariance of their noise terms (the stochastic forcing). Examining the time-varying properties of the eigenmodes with highest excitation may help elucidate the dynamical structure of the system as it transitions through a seizure.

The excitations, characteristic frequencies and damping times of the $16 \times 7 = 112$ eigenmodes of the fitted VAR[7] model were obtained for each time window as described in Section 2.2.3. Each eigenmode coefficient sequence was characterized as a relaxator or oscillator based on the sign and/or reality of the associated eigenvalue. The eigenmodes were sorted for each window by their dynamical importance. The 9 (8%) most frequently dominant eigenmodes were selected for further analysis. Figure 8 shows the damping times (left) and characteristic frequencies (right) of the dominant eigenmodes. The blue dots show the estimated damping time or frequency for each window. The estimated quantities were also smoothed with a weighted least-squares (lowess) regression using a span of 20 points (with 1-second step size for VAR estimation, this yeilds a time span of 21 seconds). This produces the curve superimposed in black. The onset (offset) of each seizure is marked by a green (magenta) vertical line. Beneath each subplot is indicated whether the eigenmode in the corresponding time window is characterized as a relaxator (blue) or oscillator (red).

Examining first the characteristic frequencies (Fig. 8-left), we see that, in the pre-ictal period, the majority of the leading eigenmodes are either characterized predominantly as relaxators or low-frequency oscillators (0.5-3 Hz). One of the eigenmodes (panel 8) appears to be an 8-12 Hz oscillator. This is characteristic of drowsy resting EEG, which is dominated by low-frequency delta (0.5-3 Hz) and alpha (8-12 Hz) oscillatory rhythms. At ictal onset we see all the leading eigenmodes dramatically shift to beta-band (12-25 Hz) oscillators with a predominance of ˜15 Hz oscillators. Note that the dominance of relaxators in the leading eigenmode (panel 1) has all but dissapeared for the entire seizure period. This indicates a transition of the global dynamics of the system to a beta-oscillatory mode. In examining the ERSP and SdDTF time-frequency images we can see that, in the early part of the seizure, the frequency dynamics and interactions are predominantly concetrated around the alpha and beta bands. Throughout the first part of the seizure, several leading eigenmodes show some slowing of the characteristic frequency (e.g., from 20 to 15 Hz). In the mid-seizure (˜270 sec for seizure 1) several leading eigenmodes show a sharp decrease in characteristic frequency to an alpha or delta-theta mode followed by a return to the orginal beta-oscillation. This is around the time when we see a strong reversal in the principal direction of information flow in the prefrontal network. Towards the end of the seizure (˜300 sec for seizure 1), for several eigenmodes, we see a brief shift in system dynamics towards low-frequency (delta-theta) oscillatory or a relaxatory mode followed by a shift in the system dynamics to an 8-10 Hz oscillatory mode, followed by a sharp decline back to a low-frequency oscillator/relaxator mode at ictal offset. This third ictal stage is approximately the time period when the information flow dynamics switch from more local interaction to strongly prefrontal $\rightarrow$ precentral interaction. The inter-ictal period following the first seizure is again dominated by infraslow (< 1 Hz) oscillatory or relaxatory dynamics, which is consistent with the surpressed neuronal state commonly observed following periods of intense ictal activity. The second seizure is similar to the first, but there does not appear to be as clear a transitioning between stages, rather the frequency dynamics appear to more smoothly slow from beta to delta throughout the seizure. The alpha mode (panel 8) seems to be the notable exception within this set of eigenmodes. Although it also exhibits a shift to the beta band during the seizure, it returns to an alpha rhythm during the inter-ictal and post-ictal periods.

Turning our attention to the damping times, we see a clear transition in system dynamics during the seizure. For the leading eigenmodes, the damping time is significantly decreased just before and during the seizure. In the inter-ictal period following the first seizure, we see a steady, dramatic, increase in damping time, followed by an abrupt decrease back to a low damping time during the second seizure. In the post-ictal period following the second seizure, we again see a steady, significant increase in damping time which does not fully decrease back to baseline within the observed interval. While the precise dynamical mechanisms underlying this phenomena are not yet fully ap-

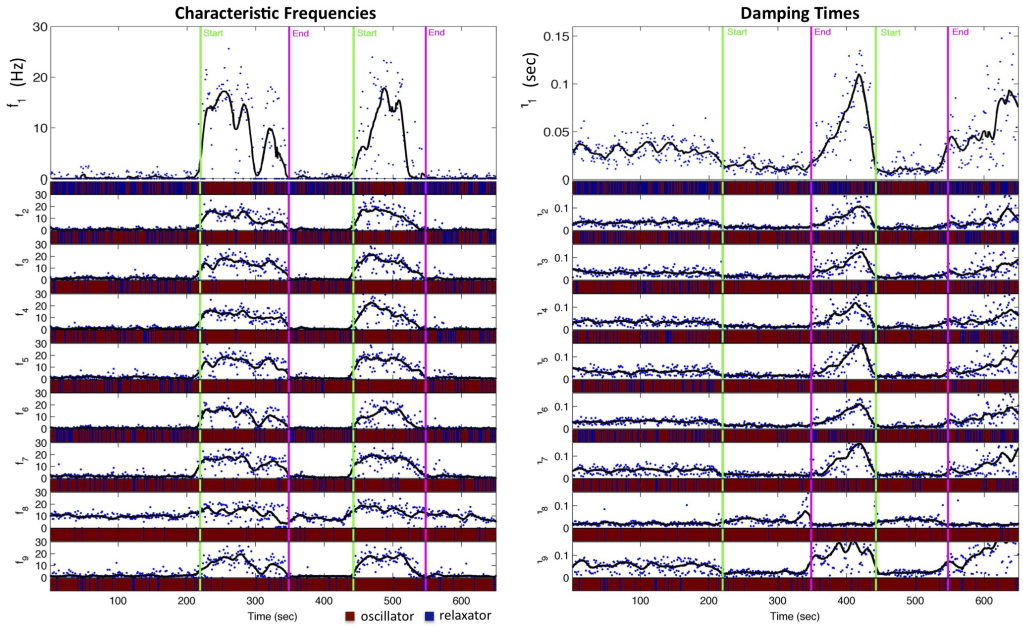

Figure 8: Characteristic frequencies (left) and damping times (right) for the 8% most dynamically important eigenmodes, in descending order of importance. Green (Magenta) vertical lines denote onset (offset) of seizure.

parent to the author, from the points mentioned in section 2.2.3, a low damping time of oscillatory eigenmodes during seizure indicates that these high-frequency phenomena are transient and likely to rapidly decay towards zero and/or may be easily perturbed by unobserved or nonlinear dynamical processes. In fact, much of the oscillatory activity during a seizure is damped, occuring in amplitude-modulated bursts lasting 1-2 seconds. Figure 9 shows such a bursting pattern for one IC. In contrast, during the inter- and post-ictal periods, the EEG is dominated by low-frequency oscillations or relaxators which may take a long time to decay to a small fraction of their original amplitude. It is worth noting that another paper which applied this eigendecomposition method to scalp EEG data, albeit using a simplified VAR[1] model, also reported a significant decrease in damping time during seizure [13]. Note that our alpha-oscillatory eigenmode (#8), exhibits an opposite pattern of activity – namely it has a low damping time during pre-, inter-, and post-ictal periods, and damping time significantly increases during seizure. Examination of other eigenmodes (not shown) revealed similar (and other complex) patterns. Future work will seek to better understand the relationship between damping time, stability, and seizure dynamics. However, a key point here is that the damping time of dominant eigenmodes is significantly modulated by ictal activity and therefore may provide a useful additional index of seizure onset.

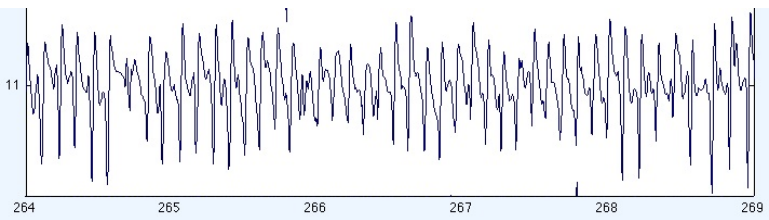

Figure 9: Bursting activity of IC11 during seizure. Note the superimposed 12 and 16 Hz damped oscillations

4 Conclusions and Future Work

In this paper we analyzed neuronal dynamics during epileptic seizures using adaptive multivariate autoregressive models applied to maximally-independent (ICA) sources of intracranial EEG data recorded from subdural electrodes implanted in a human patient for presurgery monitoring. We analyzed the time-frequency dynamics of directed information flow between sources using a multivariate granger-causal method. Seizure propagation appeared to be primarily maintained in the alpha and beta bands (with prominent peaks at 12, 15, and 20 Hz). We observed distinct stages of alternating feedforward and feedback information flow between proximal gyral and sulcul sources in a prefrontal network (elements of which also appeared to be the primary epileptogenic focus). This activity may have been maintained through short U-fiber connections. This was followed in a final stage of the seizure by a strong asymmetric spread of sustained alpha-beta ictal activity from this anterior frontal network to a posterior frontal (precentral gyrus) network, possibly though cortico-cortical white matter tracts or subcortical U-fibers. We then further examined the spatial distribution in the source domain of causal outflow and inflow using a novel combination of causal flow metrics and SBL-based source localization. This extended the previous analysis, revealing causal source and sink hubs emerging during different stages of the seizure. To our knowledge, this represents the first time these approaches have been combined to analyze spatially-localized information flow dynamics in epilepsy (or any other electrophysiological data).

To better understand the neural dynamics during the seizure, we applied an eigendecomposition method to decompose the 16-dimensional adaptive VAR[7] model into $16 \times 7 = 117$ oscillators and relaxators with characteristic damping times and frequencies. Analysis of the time-varying characteristics of the 9 (8%) most dynamically important eigenmodes revealed a prominent shift in the global state of the system from relaxatory, or low-frequency oscillatory, dynamics with a moderate damping time to beta oscillatory dynamics with low damping time at seizure onset and throughout each seizure. The distinct differences in eigenmode dynamics before, during, and after seizure suggests that analysis of characteristic frequencies and damping times of the most dynamically important eigenmodes may provide a means for detecting seizure onset. Successful intervention in a seizure (e.g., through microstimulation, TMS, or pharmacological means) requires, not only knowing where in the brain to critically intervene (which could be acheived through a combination of our proposed multivariate causal hub analysis and other traditional univariate spectral methods), but also whether it is highly likely that a seizure is occuring or about to occur. Thus a combination of the above eigenmode and causal flow and spectral analyses may provide a novel means for acheiving both goals under a unified adaptive VAR[p] modeling approach.

This paper represents a preliminary investigation into the topic. The temporal resolution and goodness-of-fit of the VAR model may be improved through the use of a Kalman filter-based adaptive VAR model (this was implemented for this paper, but due to time constraints in refining the model, results for the segmentation-based AMVAR were reported). Statistical significance will need to be assessed using phase randomization techniques. Future work will seek to further elucidate the dynamical structure of epileptogenic neuronal sources through examination of the contributions of different eigenmodes to each source. It can also be shown that the eigendecomposition provides a natural means for obtaining the multivariate spectrum of the process, which could provide alternate means for examining synchronization and coupling between sources [20]. We also plan to further examine the contributions of different frequency components to local and long-range feedforward and feedback influences in seizure propagation.

Acknowledgments

A **huge** thanks to Zeynep Akalin Acar for providing SBL source localization for the ICs and helping immensely in creating the causal flow movies. Thanks also to Greg Worrell (Mayo Clinic, Rochester, MN) for providing the data. This paper used modified routines from the ARFIT [19] software package as well as the author's SIFT software package [16].

References

[1] Z. Acar and G. Worrell. Patch-based cortical source imaging in epilepsy. In *IEEE EMBC*, 2009.

[2] Z. A. Acar, S. Makeig, and G. Worrell. Head modeling and cortical source localization in epilepsy. In *IEEE EMBC*, volume 2008, pages 3763–6, Jan. 2008.

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

[4] S. L. Bressler and A. K. Seth. Wiener-Granger Causality: A well established methodology. *NeuroImage*, 2010.

[5] E. Bullmore and O. Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature reviews. Neuroscience*, 10(3):186–98, 2009.

[6] J. Burg. Maximum entropy spectral analysis. In *37th Ann. Int. Meet., Soc. Explor.Geophys.*, Oklahoma City, OK, USA, 1967.

[7] J. Burg. *Maximum entropy spectral analysis*. Stanford University Press, Stanford, CA, USA, 1975.

[8] G. Buzsaki. *Rhythms of the Brain*. Oxford University Press, USA, 2006.

[9] G. Florian and G. Pfurtscheller. Dynamic spectral analysis of event-related EEG data. *Electroencephalography and clinical neurophysiology*, 95(5):393–396, Nov. 1995.

[10] M. Ge, X. Jiang, Q. Bai, S. Yang, J. Gusphyl, and W. Yan. Application of the directed transfer function method to the study of the propagation of epilepsy neural information. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 2007:3266–9, Jan. 2007.

[11] B. H. Jansen, J. R. Bourne, and J. W. Ward. Autoregressive estimation of short segment spectra for computerized EEG analysis. *IEEE transactions on bio-medical engineering*, 28(9):630–8, Sept. 1981.

[12] T. W. Lee, M. Girolami, and T. J. Sejnowski. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural computation*, 11(2):417–41, Feb. 1999.

[13] X. Li, X. Guan, and R. Du. Using Damping Time for Epileptic Seizures Detection in EEG. *Time*, (x):1–4, 2007.

[14] H. Lütkepohl. *New Introduction to Multiple Time Series Analysis*. Springer, Berlin, Germany, 2006.

[15] S. Makeig, A. J. Bell, T.-p. Jung, and T. J. Sejnowski. Independent component analysis of electroencephalographic data. *In D. Touretzky, M. Mozer, and M. Hasselmo, editors, Advances in Neural Information Processing Systems*, 8pages:145–151, 1996.

[16] T. Mullen, A. Delorme, C. Kothe, and S. Makeig. An Electrophysiological Information Flow Toolbox for EEGLAB. In *Society for Neuroscience*, San Diego, CA, 2010.

[17] A. Neumaier and T. Schneider. Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Transactions on Mathematical Software (TOMS)*, 27(1):27–57, 2001.

[18] J. Schleimer. *Phase Synchronisation in Superimposed Electrophysiological Data*. PhD thesis, University of Helsinki, 2007.

[19] T. Schneider and A. Neumaier. Algorithm 808: ARfit—a matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Transactions on Mathematical Software*, 27(1):58–65, Mar. 2001.

[20] H. von Storch, G. Burger, R. Schnur, and J. von Storch. Principal Oscillation Patterns: A Review. *Journal of Climate*, 8(3):377–400, 1995.

[21] C. Wilke, L. Ding, and B. He. An adaptive directed transfer function approach for detecting dynamic causal interactions. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 2007(2):4949–52, Jan. 2007.

[22] C. Wilke, L. Ding, and B. He. Estimation of time-varying connectivity patterns through the use of an adaptive directed transfer function. *Biomedical Engineering, IEEE Transactions on*, 55(11):2557–2564, 2008.

[23] D. Wipf, R.R. Ramirez, J.A. Palmer, and S. Makeig, and B.D. Rao. Analysis of empirical Bayesian methods for neuroelectromagnetic source localization. *NIPS*, 1505-1512, 2007.

[24] L. Zetterberg. Estimation of Parameters for a Linear Difference Equation with Application to EEG Analysis. *Mathematical Biosciences*, 5(3-4):227–275, 1969.