

Delay Differential Analysis of Seizures in Multichannel Electrocorticography Data

Claudia Lainscsek

claudia@salk.edu

Howard Hughes Medical Institute, Computational Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037, U.S.A., and Institute for Neural Computation, University of California San Diego, La Jolla, CA 92093, U.S.A.

Jonathan Weyhenmeyer

jweyhenm@iupui.edu

Goodman Campbell Brain and Spine, Department of Neurological Surgery, Indiana University School of Medicine, Indianapolis, IN 46202, U.S.A.

Sydney S. Cash

scash@mgh.harvard.edu

Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, U.S.A.

Terrence J. Sejnowski

terry@salk.edu

Howard Hughes Medical Institute, Computational Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037, U.S.A., and Institute for Neural Computation, University of California San Diego, La Jolla, CA 92093, U.S.A.

High-density electrocorticogram (ECoG) electrodes are capable of recording neurophysiological data with high temporal resolution with wide spatial coverage. These recordings are a window to understanding how the human brain processes information and subsequently behaves in healthy and pathologic states. Here, we describe and implement delay differential analysis (DDA) for the characterization of ECoG data obtained from human patients with intractable epilepsy. DDA is a time-domain analysis framework based on embedding theory in nonlinear dynamics that reveals the nonlinear invariant properties of an unknown dynamical system. The DDA embedding serves as a low-dimensional nonlinear dynamical basis onto which the data are mapped. This greatly reduces the risk of overfitting and improves the method's ability to fit classes of data. Since the basis is built on the dynamical structure of the data, preprocessing of the data (e.g., filtering) is not necessary. We performed a large-scale search for a DDA model that best fit ECoG recordings using a genetic algorithm to qualitatively discriminate between

different cortical states and epileptic events for a set of 13 patients. A single DDA model with only three polynomial terms was identified. Singular value decomposition across the feature space of the model revealed both global and local dynamics that could differentiate electrographic and electroclinical seizures and provided insights into highly localized seizure onsets and diffuse seizure terminations. Other common ECoG features such as interictal periods, artifacts, and exogenous stimuli were also analyzed with DDA. This novel framework for signal processing of seizure information demonstrates an ability to reveal unique characteristics of the underlying dynamics of the seizure and may be useful in better understanding, detecting, and maybe even predicting seizures.

1 Introduction

Differences in transmembrane currents in separate neuronal and cortical areas give rise to electric fields that can be measured macroscopically (Buzsaki, Anastassiou, & Koch, 2012). Macroscopic cortical measurements span many orders of magnitude; electrocorticogram (ECoG) electrodes, for example, record from a much smaller cortical area than scalp electroencephalogram (EEG) electrodes (Nunez & Srinivasan, 2006). Increasing the spatial scale of recording, and thereby decreasing the spatial resolution, involves a natural trade-off of information loss in the form of data reduction via spatial averaging (Nunez & Srinivasan, 2006). On smaller scales, local field potentials (LFPs) contain more localized spatial information than ECoG and EEG recordings.

A significant portion of neurophysiological analysis spanning the entire cortical recording spectrum has focused on different forms of frequency analysis, ranging from power confined to spectral bands to coherence within and between bands across and within neuronal populations (Uhlhaas & Singer, 2006). It is therefore entirely plausible that the analysis of EEG and ECoG data simultaneously acquired from the same experiment may yield conflicting results. For example, epilepsy was originally assumed to be the result of highly synchronous global activity (Jiruska et al., 2012); however, recent research has shown that at different scales, the population activity is not only asynchronous (Netoff & Schiff, 2002) but such activity may act as a primer for the epileptic state (Dominguez et al., 2005). Thus, it would be beneficial to find a method that is able to characterize the internal state of the cortex rather than characterizing the outputs of the system. In other words, a model that reflected how the internal states are generated would be more revealing than a model of the recordings, which reflect additional generators, such as artifacts, that obscure the internal states.

Epilepsy has been hypothesized to be caused by an underlying deterministic, nonlinear system that undergoes phase transitions between pre-ictal, ictal, and post-ictal states (Iasemidis & Sackellares, 1996; Kramer et al.,

2013). Other studies have posited that deterministic, nonlinear dynamics become more profound during ictal activity (Andrzejak et al., 2001). Consequently, much emphasis has been placed on attempting to model epileptiform phase transitions as bifurcations that bridge chaotic and simple states (Kramer et al., 2013; da Silva et al., 2003; Velazquez, Cortez, Snead, & Wennberg, 2003). Nonlinear measures such as correlation dimension (Elger & Lehnertz, 1998; Lehnertz & Elger, 1998), Lyapunov exponents (Iasemidis & Sackellares, 1996), and entropy (Lehnertz et al., 2003) have been postulated to reliably predict the onset of ictal activity in patients with epilepsy. Explicitly, when the brain moves from a non-ictal state to an ictal state, it loses complexity, reduces dimensionality, and becomes more predictable. As such, the epileptic EEG can be modeled as a low-dimensional system where global features can be used to characterize and classify the underlying dynamical system (Quyen, Martinerie, Adam, & Varela, 1997; Pijn et al., 1997). The major drawback to traditional nonlinear measures (e.g., correlation dimension and Lyapunov exponent) is that they are average measures requiring a large number of samples and have significant problems with nonstationary signals (Pijn et al., 1997; Hornero, Abasolo, Escudero, & Gomez, 2009). Indeed, epileptic events are inherently nonstationary and undergo fast phase transitions, requiring an analysis technique that is robust to short data segments and nonstationarity. Moreover, many of the noted techniques were initially tested on limited data sets, and additional study on long-term recordings has not yielded promising results (Mormann, Andrzejak, Eger, & Lehnertz, 2007; Stacey, Quyen, Mormann, & Schulze-Bonhage, 2011) with respect to seizure prediction.

This article is organized as follows. Section 2 provides patient information and introduces the methods. In section 3, DDA results are shown for clinical and nonclinical seizures, stimulations, artifacts, and interictal data. Section 4 compares clinical findings with DDA results and shows a comparison with spectral features. Section 5 is the discussion, and section 6 is the conclusion. Sections 2.3, 2.5, and 2.6 are mainly targeted to readers interested in the details of the method.

1.1 Delay Differential Analysis and Delay Differential Equations.

Extracting relevant features from ECoG data requires a method that is capable of high temporal resolution and dimensionality reduction. Dimensionality reduction is important because not all possible states of a biological system are needed to characterize normal or pathological conditions. Methods such as principal component analysis (PCA), independent component analysis (ICA), and frequency analysis are often used to reduce the dimensionality of a given system; however, these linear techniques tend to fail when they are applied to systems that are inherently nonlinear because they describe the system in Euclidean space rather than in coordinates consistent with the system's true underlying manifold (Tenenbaum, Silva, & Langford, 2000). Excessive reduction may also fail to identify relevant features

because false nearest neighbors may occur in embedding space when the dimensionality of the embedding is too small, resulting in the false projection of nonneighboring points into a space where the distance between the points is smaller than the distance between them in a space of adequate dimensionality. Thus, the reduction in dimensionality must be balanced to properly unfold the dynamics of the system while avoiding a false representation.

Delay differential analysis (DDA) is a functional embedding technique based on delay differential equations (DDEs) that provides a highly specific, low-dimensional classification of an observable time series. DDEs are composed of two embeddings: a delay embedding and a derivative embedding. An embedding is the mapping of a single time series into a multidimensional object in embedding space (Whitney, 1936; Packard, Crutchfield, Farmer, & Shaw, 1980; Takens, 1981; Sauer, Yorke, & Casdagli, 1991). The delay embedding is composed of a time series $x(t)$ and its delayed versions $x(t - \tau_n)$, where the delays τ_n are restricted to positive integer multiples of the time steps δt . DDEs are considered to be nonuniform because no restriction is placed on the relationship between delays found in a given model; for example, τ_j is not required to be a multiple of any other time delay τ_i found in the model (Judd & Mees, 1995). A derivative embedding is composed of a single time series $x(t)$ and its n -order derivatives. Thus, an embedding converts a single time series into a multidimensional geometrical object that is well suited for classification and topographical analysis.

DDEs capture essential features in data to produce classification outputs without having direct access to all of the system's variables (Kremlivsky & Kadtke, 1997; Kadtke & Kremlivsky, 1997; Lainscsek, Weyhenmeyer et al., 2013; Lainscsek & Sejnowski, 2015). This property has allowed for DDE-based algorithms to robustly discriminate pathological and healthy EEG recordings from subjects for a wide range of disease and recording paradigms (Lainscsek et al., 2012; Lainscsek, Hernandez et al., 2013; Lainscsek & Sejnowski, 2013). Additionally, analysis of single-electrode EEG data has proven the ability of DDEs to uniquely identify all sleep stages as well as discriminate between rapid eye movement (REM) and awake states (Lainscsek, Messager, Portman, Sejnowski, & Letellier, 2014).

The classification performance on data that are inseparable by linear methods (e.g., frequency analysis) is a direct result of two overarching properties of nonlinear dynamics. First, it is possible to reconstruct a dynamical system without access to all representative time series of that system (Takens, 1981). Second, if data are considered to be linearly inseparable in Euclidean space, it is possible to find a nonlinear transformation that produces a desired linear separation between two classes of data provided the nonlinear transformation is of adequate dimensionality (Cover, 1965). Thus, provided there are classes of data generated by different underlying dynamical systems or a single dynamical system in different dynamical states, it is possible to linearly separate the classes in a nonlinear space.

Indeed, DDEs are able to identify different states in a single dynamical system because of the relationship between the timescales and the dynamical parameters of the system. The ability to capture small changes in both timescales and dynamics arises from the nonuniform embedding property relating the time series $x(t)$ and its delayed version $x(t - \tau_n)$ to its derivative $\dot{x}(t)$. DDEs that are matched to the dynamics of a system such as the cortex are capable of robustly discriminating between different states within the same dynamical system. Such discrimination is immediately applicable to neuronal population activity as measured by ECoG and related modalities.

Here, we focus on the application of DDA to ECoG obtained from a group of 13 patients with medication-refractory epilepsy. ECoG signals obtained from two patients were the focus of the analysis we present here. Our goal is to establish that DDA is a novel and useful time-domain technique for analyzing ECoG data and one that shows promise for identifying distinct ECoG events. In subsequent studies, we will use the features extracted from these events to characterize the underlying cortical states of a larger population of patients with epilepsy.

2 Materials and Methods

2.1 Patient Information and Data Recordings. Electrocorticography from 155 seizures (2 to 74 seizures per patient; mean of 11.9) in 13 patients with long-standing pharmaco-resistant complex partial seizures were analyzed. Multiple etiologies were represented in this patient population, including mesial temporal sclerosis (MTS) as diagnosed by history, seizure semiology, and imaging; cortical dysplastic lesions confirmed with pathology; and other etiologies. Patient recordings were selected that were known to contain seizures with focal onset with and without typical complex partial events with secondary generalization. Determination of the seizure onset time was made by observing when intracranial electrodes began to show standard ictal electrographic activity, including low-voltage fast activity or repetitive spike-wave discharges. All recordings were performed using a standard clinical recording system (XLTEK, Natus Medical, San Carlos, CA) with a 500 Hz sampling rate. The reference channel was a strip of electrodes placed outside the dura and facing the skull at a region removed from the main grid of electrodes. This was most often over posterior parietal regions. Subdural electrode arrays were placed to confirm the hypothesized seizure focus and locate epileptogenic tissue in relation to essential cortex, thus directing surgical treatment. The decision to implant, the electrode targets, and the duration of implantation were made entirely on clinical grounds with no input from this research study. All data analyses were performed under protocols monitored by the Massachusetts General Hospital according to National Institutes of Health guidelines.

2.2 Delay Differential Analysis. Henceforth, we will consider all DDEs of interest to be of the general polynomial form,

$$\dot{x} = \sum_{i=1}^I a_i \prod_{n=1}^N x_{\tau_n}^{m_{n,i}} \quad \text{for } \tau_n, m_{n,i} \in \mathbb{N}_0 \quad (2.1)$$

with N delays, I monomials with coefficients a_1, a_2, \dots, a_I , and a degree $\max(m_{n,i})$ of nonlinearity where $x = x(t)$ and $x_{\tau_n} = x(t - \tau_n)$. Any monomials that are not included will have the corresponding a_i set to 0. We restricted the model space to models composed of two delays ($N = 2$) and up to fourth-order nonlinearity ($\sum_n m_{n,i} \leq 4$), yielding a list of 14 possible monomials. The most general 14-term DDE is of the form

$$\begin{aligned} \dot{x} &= f(a_i, x_{\tau_n}) \\ &= a_1 x_{\tau_1} + a_2 x_{\tau_2} + a_3 x_{\tau_1}^2 + a_4 x_{\tau_1} x_{\tau_2} + a_5 x_{\tau_2}^2 + a_6 x_{\tau_1}^3 \\ &\quad + a_7 x_{\tau_1}^2 x_{\tau_2} + a_8 x_{\tau_1} x_{\tau_2}^2 + a_9 x_{\tau_2}^3 + a_{10} x_{\tau_1}^4 + a_{11} x_{\tau_1}^3 x_{\tau_2} \\ &\quad + a_{12} x_{\tau_1}^2 x_{\tau_2}^2 + a_{13} x_{\tau_1} x_{\tau_2}^3 + a_{14} x_{\tau_2}^4. \end{aligned} \quad (2.2)$$

Only models consisting of one, two, or three monomial terms ($I \leq 3$) were considered in our analysis. The coefficients a_i were estimated with numerical singular value decomposition (SVD) to minimize the least square error (Press, Flannery, Teukolsky, & Vetterling, 1990). The value of each time delay, τ_n , found within a given model was allowed a range between 1 and $200 \delta t$ where $\delta t = \frac{1}{f_s} = 2$ ms, $f_s = 500$ Hz. The derivative of a single short time series was computed using a five-point center derivative algorithm (Miletics & Molnárka, 2004). Note that the DDE models $\dot{x} = a_1 x_{\tau_1} + a_4 x_{\tau_1} x_{\tau_2}$ and $\dot{x} = a_2 x_{\tau_2} + a_4 x_{\tau_1} x_{\tau_2}$ are the same with exchanged delays τ_1 and τ_2 . Therefore only the first of these two models was used. All such redundant DDE models were omitted, allowing for a model space reduction from 469 to 245.

In general, $\dot{x}(t) - f(a_i, x_{\tau_n}) \neq 0$ (i.e., the summation of the time series $x(t)$ and its delayed forms x_{τ_n}) is not guaranteed absolute equality with $\dot{x}(t)$. Indeed, there is an inherent error when fitting any system with models that are bounded with respect to the number of terms, order of nonlinearity, and permitted delay values. The deviation between model delay pair $f(a_i, x_{\tau_n})$ and signal derivative $\dot{x}(t)$ is calculated using least square error,

$$\rho = \sqrt{\frac{1}{K} \sum_{k=1}^K \left(\dot{x}_{t_k} - \sum_{i=1}^I a_i \prod_{n=1}^N x_{t_k, \tau_n}^{m_{n,i}} \right)^2}, \quad (2.3)$$

for K data points.

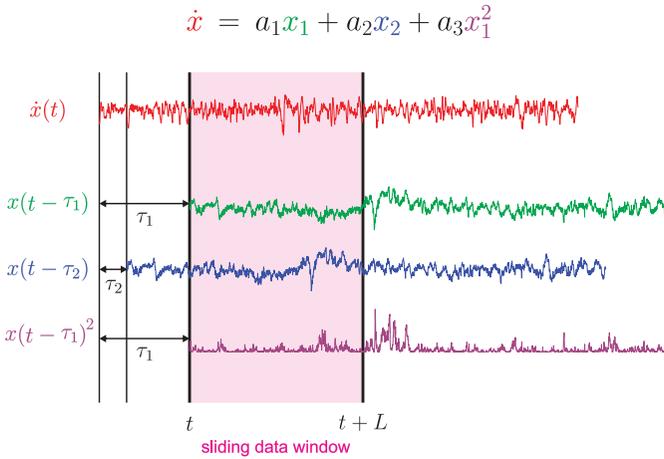


Figure 1: Estimation of the DDA features using sliding windows.

Figure 1 illustrates how the DDA features are computed from the data on a three-term DDE. After normalizing each data window to zero mean and unit variance, the coefficients a_i are estimated using SVD, and the error is computed with equation 2.3.

The overarching objective of this analysis is to identify groups of data based on timescale and dynamical characteristics. We do not attempt to model a particular time series or data segment. Restrictions placed on model terms, order, and delay values force the selected model-delay parameter space to fit the dynamics of classes rather than those of a particular event in a particular time series (e.g., overfitting). Restricting the number of monomial terms greatly decreases the probability that the selected DDE is modeling noise and increases the likelihood that it instead captures specific properties of the dynamical system. The error, ρ , and coefficients, $a_1, \dots, a_i, i \leq 3$, are considered discriminative features that are used to differentiate between different dynamical systems and system states. The dimensionality $d \leq 4$ is low compared with linear spectral feature analysis.

2.3 Structure Selection

2.3.1 Evolutionary Computation of Models and Time Delays. Genetic algorithms (GA) are a population-based search method whereby a random initial population of parameters is manipulated by a set of operators that cause both small and large changes to the individual members of the population. The postoperator population is then tested against a fitness function that effectively details how well a parameter set solves a predefined problem. The GA described here chooses model and delay pairs that discriminate

between specific states of ECoG data, and the fitness function was defined to be the least square error, equation 2.3, obtained from the comparison of model and delay pair and the signal derivative.

In accordance with basic genetic algorithms (Goldberg, 1998), the present GA used the natural operators of reproduction, mutation, and crossover. In order to simplify the implementation of the GA, all of the information with respect to the model and delay parameter space was encoded into binary strings (Goldberg, 1998). Because the DDE model and delay pairs were constrained (e.g., models were limited to one to three terms, two delays permitted to range from $1 - 200 \delta t$, and fourth-order nonlinearity), 24 bits were required such that no duplicate representations existed for nonidentical model-delay pairs. Each model and each delay was encoded with 8 bits. The complete binary string representations of model and delay pairs are called chromosomes. Models and delays were given specific, nonvarying locations on the chromosome; for example, the model was encoded in bits 1 to 8, the first delay was encoded in bits 9 to 16, and the second delay was encoded in bits 17 to 24. Both the delay and model ranges failed to fully use all possible binary representations given by their bit allotment. Any remaining binary representations were filled with random delays and models in order to avoid creating a large, unusable subspace of the solution space. This allocation did not greatly bias our results as a model or delay is likely to be represented at most twice within a particular space and the increased representation is unlikely to continue for successive initializations (e.g., the extra models and delays will change from one data window to the next).

The fitness, reproduction, mutation, and crossover operators are displayed in Figure 2. A chromosome's fitness is defined by how well the specific model and delay pair fit a particular segment of data (e.g., a minimization of the mean square error). Thus, fitter chromosomes have a smaller mean square error. Reproduction is defined as the act of replicating an entire chromosome for the next generation of a population. Reproduction is performed in a manner such that fitter chromosomes make up a larger portion of the progeny population. Specifically, the fittest individual is replicated the most times, and a decreasing monotonic function dependent on a model delay pair's corresponding ρ ,

$$m_i(t + 1) \propto f(\rho_i)m_i(t), \quad (2.4)$$

describes each model and delay pair's proportional representation in the progeny population. Note that $f(\rho_i)$ is a function that increases with decreasing ρ_i and m_i is a specific member of the population. Mutation is the act of flipping a bit within a binary string (e.g., a 1 becomes a 0). Importantly, each bit has equal probability of being flipped. Because the delays are a direct encoding from integer to binary, the integer value assigned to each binary bit increases when moving leftward in a string. As such, larger

deviations in integer values are obtained when mutating more leftward bits. Indeed, large mutations are as probable as small mutations, and the algorithm is equally likely to sample a large or small neighborhood of a model-delay string undergoing mutation. Models were encoded as a binary index that ranged from 1 to 256, and any mutation was likely to greatly change the model. The mutation rate was varied between 0.1 and 0.4 for each mutation cycle. Crossover is explicitly defined as the act of trading components of the model delay pairs. Specifically, the points of separation between model, first delay, and second delay are crossover points about which one chromosome may switch components with another chromosome. Crossover was not allowed within component strings.

The GA was implemented with additional tools to perform robust searches based on best-fit delay parameters within the present population. Here, we assume that a well-performing delay captures an intrinsic property of the signal in question and may provide better fit when it is applied to other model terms. For each population, the model-delay pairs were ranked according to fitness. A subset of best delay pairs is then used to create a mutated population,

$$\boldsymbol{\tau}_{search} = \begin{pmatrix} \boldsymbol{\tau}_{best} - 1 \\ \boldsymbol{\tau}_{best} \\ \boldsymbol{\tau}_{best} + 1 \end{pmatrix}, \quad (2.5)$$

where $\boldsymbol{\tau}_{best}$ is a delay array containing delay pairs from the model-delay members with the lowest error. All possible permutations of the $\boldsymbol{\tau}_{search}$ matrix were paired with the best-performing model and passed through the fitness function. The newly created model-delay pairs were added to the parameter population, and the population was sorted according to the lowest error. Only a set number of population members were retained according to the predefined maximum population size.

2.3.2 GA Performance on Simulated Data. Prior to implementing the GA on real data, we first applied the GA to a difficult test case based on DDE characteristics previously described in Lainscsek and Sejnowski (2013, 2015). Consider a linear DDE of the form

$$\dot{x} = \sum_{i=0}^{2N} a_i x_{\tau_i}. \quad (2.6)$$

It can be shown that a special solution of equation 2.6 is

$$x(t) = \sum_{i=0}^{N-1} \cos(\omega_i t); \tau_i = \frac{\pi(2n-1)}{2\omega_j} \quad (2.7)$$

with

$$\begin{aligned}
 a_1 &= -\frac{1}{4} \sec^2\left(\frac{\pi\omega_2}{2\omega_1}\right) \left(\omega_1 - \cos\left(\frac{2\pi\omega_2}{\omega_1}\right) \csc\left(\frac{\pi\omega_2}{2\omega_1}\right) \omega_2\right), \\
 a_2 &= \frac{1}{2} \sec^2\left(\frac{\pi\omega_2}{2\omega_1}\right) \left(\cos\left(\frac{\pi\omega_2}{\omega_1}\right) \omega_1 - \sin\left(\frac{3\pi\omega_2}{2\omega_1}\right) \omega_2\right), \\
 a_3 &= -\frac{1}{4} \csc\left(\frac{\pi\omega_2}{2\omega_1}\right) \sec^2\left(\frac{\pi\omega_2}{2\omega_1}\right) \left(\sin\left(\frac{\pi\omega_2}{2\omega_1}\right) \omega_1 + \cos\left(\frac{\pi\omega_2}{\omega_1}\right) \omega_2\right),
 \end{aligned}
 \tag{2.8}$$

where all τ_i are inversely related to a single frequency ω_j found in the signal and the coefficients are related to all of the frequencies. Certain τ_i are related to τ_j via odd integer multiples, and there exists an infinite number of τ_i for each frequency in the signal due to naturally occurring harmonics. Limiting the τ range allows for a test case in which a small subset τ_i can be considered the solution. In this experiment, a signal $x(t) = \cos(2\pi f_1 t) + \cos(2\pi f_2 t)$ where $f_1 = 100$ Hz, $f_2 = 125$ Hz, and $f_s = 10,000$ Hz was generated (see Figure 3A). The model $\dot{x} = a_1 x_{\tau_1} + a_2 x_{\tau_2} + a_3 x_{\tau_3}$ was fixed in order to make the exact solution less degenerate. The exact solution for the signal $x(t)$, a three-element τ vector that is a subset of the set $\{20, 25, 60, 75, 100, 125, 140, 175, 180\}$, is shown in Table 1. All τ_i in a solution vector τ relate to the same frequency, and there are no exact solutions with τ_i corresponding to both frequencies found in the signal. The solution space is graphically displayed in Figure 3B with the delay combinations indexed linearly along the ordinate axis so as to map the four-dimensional space to two dimensions. The solution space appears fractal rather than smooth, implying that a gradient descent algorithm will not adequately search the space as there are many local minima that do not lead to global minima or exact solutions. Any solution below 10^{-15} was within the precision error of zero and was considered to be an exact solution. The behavior of the GA was further characterized for populations of 100, 200, 300, 400, 500, 750, and 1000 members by running each population size 10,000 times on the synthetic data. The percentage of correct outputs for each population is shown in Figure 4. The asymptotic behavior of the correct-response curve implies that there is a population for which the trade-off between computation time and correct response rate is satisfactory. Finding such a population for real data is difficult, and instead the GA was run on populations that permitted convergence—for example, populations that provide the same output when the same data segment is analyzed multiple times.

2.3.3 GA Analysis of ECoG Data. For the structure selection analysis on ECoG data, we used short 2 s data windows with no overlap between successive windows. Data segments were sampled at 500 Hz, yielding data

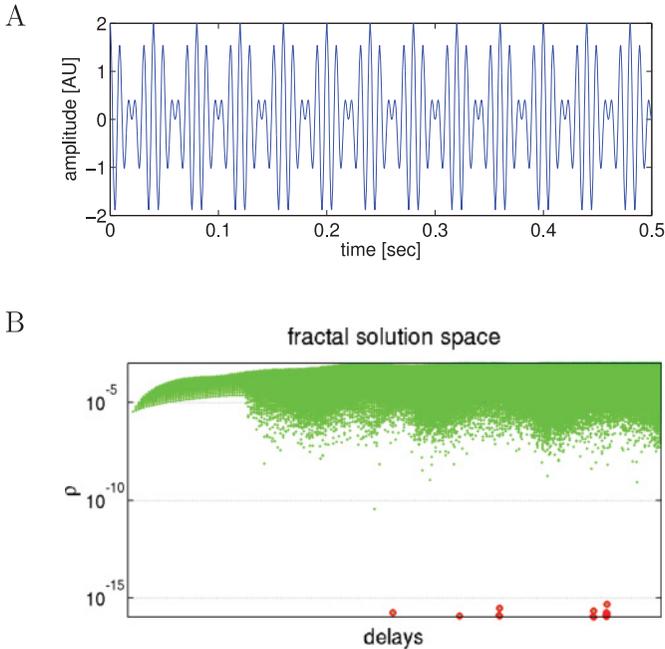


Figure 3: Test case for GA performance: (A) A synthetic signal consisting of two cosine waveforms at two different frequencies. (B) Example of delay solution space for the signal shown in the top panel. The red dots indicate an exact solution, and the green dots indicate all incorrect or inexact solutions.

Table 1: List of Possible Solutions for Cosine Waveform Shown in Figure 3A.

Number	τ_1	τ_2	τ_3	#	τ_1	τ_2	τ_3
1	20	60	100	8	75	125	175
2	25	75	125	9	20	60	180
3	20	60	140	10	20	100	180
4	20	100	140	11	60	100	180
5	60	100	140	12	20	140	180
6	25	75	175	13	60	140	180
7	25	125	175	14	100	140	180

windows that contained 1000 samples. A model-delay pair search was conducted with a GA on 1 hour recordings surrounding an epileptic event. Explicitly, the GA was run with 2 s nonoverlapping windows on a data segment consisting of 1/2 hour prior to clinically defined seizure onset and 1/2 hour postseizure onset. Clinically defined seizure onset channels and other randomly selected channels were included in the GA search. The

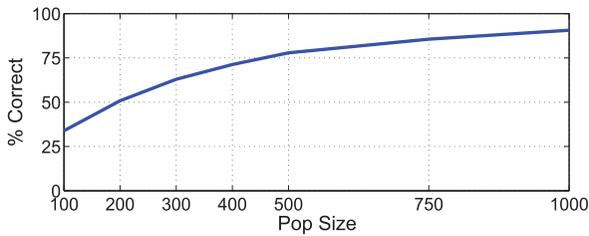


Figure 4: Relationship between population size and likelihood of the GA finding the correct solution on synthetic data.

choice of nonoverlapping 2 s windows permitted a general characterization of the data set such that the model-delay pairs that best characterized the ECoG data in vastly different states could be ascertained. The 1 hour recordings contained many events common to ECoG analysis (e.g., artifacts, exogenous stimuli, and interictal spikes), allowing for further characterization of delay differential analysis in realistic ECoG recording sessions. The method's performance on data segments containing artifacts and external stimuli is important because no preprocessing (e.g., filtering or artifact rejection) was applied to the recordings prior to performing the delay differential analysis.

2.4 DDA versus Traditional ECoG Analysis. In traditional ECoG analysis (see Figure 5, right panel), the data need intensive preprocessing, including extensive filtering, artifact rejection, and visual inspection, prior to signal analysis. This process is not automatic and is subjective, allowing for potential interrater variability. Additionally, nonlinear information may be inadvertently removed from the data. Traditionally, spectral analysis is the next step in the processing pipeline, producing a high-dimensional spectral feature space. This can be reduced using techniques for dimensionality reduction and machine learning to identify a subset of features that provide signal characterization.

In DDA (see Figure 5, left panel), data preprocessing is reduced to amplitude removal of the raw data by normalizing to zero mean and unit variance in each data window. There is no filtering, visual inspection, or artifact rejection. The DDA processing pipeline is divided into two parts: structure selection and data analysis. Structure selection must be performed one time for each data class (e.g., ECoG, EEG, or EKG data) in order to identify a global DDE model that captures essential dynamical information of the data. The selection of the corresponding delays is task specific (e.g., data-specific event characterization). It should be noted that low-dimensional DDE models were chosen for detection, classification, and characterization of the data but not for prediction or modeling. After the

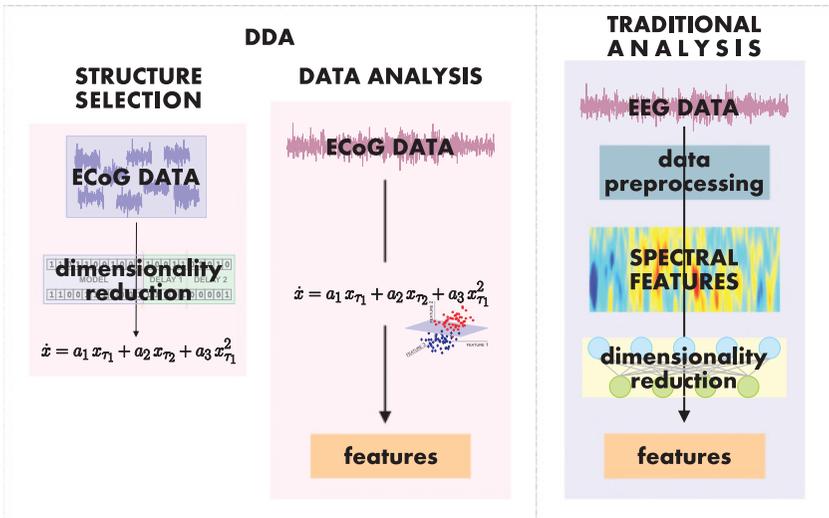


Figure 5: Data pipeline analysis for traditional ECoG analysis (right) compared with DDA (left).

DDE model is identified, data analysis is performed as a fully automatic one-step process.

2.5 Singular Value Analysis of the DDA Feature Space

2.5.1 *Singular Value Decomposition.* The decomposition of a matrix \mathbf{M} to obtain its singular values (SV) is a change of basis transformation that rotates the system’s coordinates by maximizing the variance in each of the new dimensions,

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \tag{2.9}$$

where $\mathbf{M} \in \mathbb{R}^{N \times n}$,

$$\mathbf{M} = \begin{pmatrix} X_{J_1, K_1} & X_{J_1, K_2} & \cdots & X_{J_1, K_n} \\ X_{J_2, K_1} & X_{J_2, K_2} & \cdots & X_{J_2, K_n} \\ \cdots & & \ddots & \\ X_{J_N, K_1} & X_{J_N, K_2} & \cdots & X_{J_N, K_n} \end{pmatrix}, \tag{2.10}$$

with rank $m \leq n < N$ whose number of rows N is greater than or equal to its number of columns n , can be written as the product of an orthogonal

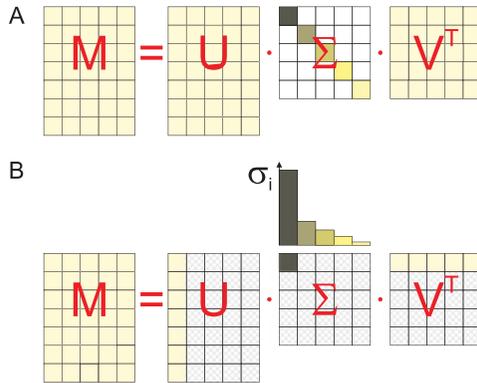


Figure 6: Schematic matrices illustrating (A) singular value decomposition of nonsquare matrix M into orthonormal matrices, U and V , and a diagonal ranked singular value matrix Σ . (B) Singular value decomposition of matrix M to retain only the most significant singular value (dark gray entry in Σ , which eliminates most of the other entries, in light gray cross-hatch).

matrix $U \in \mathbb{R}^{N \times N}$, a diagonal matrix $\Sigma \in \mathbb{R}^{N \times n}$ with positive or zero elements $s_1 \geq s_2 \geq \dots > s_m > s_{m+1} = \dots = s_n = 0$ (the singular values), and the transpose of an orthogonal matrix $V \in \mathbb{R}^{n \times n}$. The variables X , J , and K will be identified and explained in section 2.5 because multiple forms of SV analysis will be introduced. Such a rotation is favorable when analyzing real data as the system’s noise is decreased relative to the number of dimensions retained in the projection (Schiff, Sauer, Kumar, & Weinstein, 2005). The result is a projection of the system into a less noisy space where essential qualitative and quantitative information about the system is retained (Broomhead & King, 1986). In this projection, large events (e.g., seizures) would be retained and smaller or spurious events would be removed, leaving a clean representation of the events of interest. A graphical explanation of SVD into unitary orthonormal matrices, U and V , and diagonal singular value matrix Σ and is shown in Figure 6A, with an example of retention of the first singular value in Figure 6B.

2.5.2 Tuncated Higher-Order Singular Value Decomposition (THOSVD) of the DDA Feature Space. As previously mentioned, ECoG data are high dimensional and rich in spatial and temporal information. Performing SVD on ECoG recordings across time and space (channels) would yield a reduced but high-dimensional time-space projection that would be difficult to interpret. Rather than develop a singular system solely from ECoG signals, we chose to develop a singular system based on the DDA features. GA structure selection identifies a single DDE model and a subset of delay pairs that

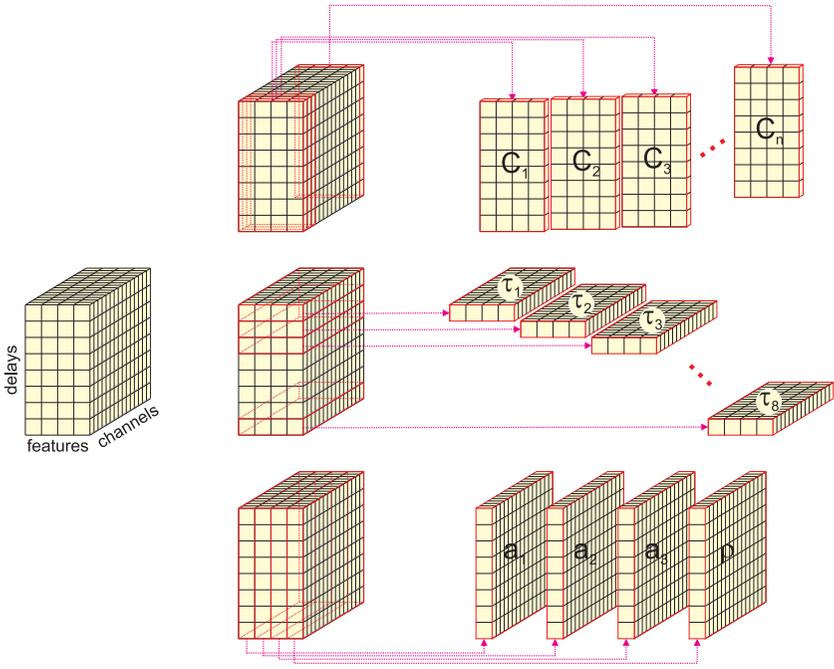


Figure 7: Example of a matrix unfolding of a three-dimensional DDA feature tensor (left) for each time window t_k where the three dimensions correspond to DDE features b_l , delay pairs τ_m , and channel numbers C_n . The three panels on the right show the tensor unfoldings into delay-feature (C_i), feature-channel (a_i), and delay-channel (τ_i) matrices.

characterize the ECoG data across subjects and events (shown in section 3). This results in a four-dimensional DDA feature tensor,

$$\mathcal{T}_{t_k, b_l, \tau_m, C_n}, \tag{2.11}$$

where the dimensions are time t_k , model features b_l ($\mathbf{b} = \{\mathbf{a}, \rho\}$), delay pairs τ_m , and channels C_n . Multilinear generalizations of SVD to multidimensional tensors have been introduced by Tucker (1966) and Lathauwer, de Moor, and Vandewalle, (2000). Here we utilize a truncated version of the multilinear SVD technique HOSVD (higher-order SVD) and call it THOSVD in the remainder of the article. Figure 7 shows a 3D example of the main step in THOSVD, the unfolding of a feature tensor to matrices. For our 4D DDA feature tensor $\mathcal{T}_{t_k, b_l, \tau_m, C_n}$, six such unfoldings to the delay-feature, feature-channel, delay-channel, time-feature, time-delay, and time-channel matrices can be made. Traditional HOSVD does a rank decomposition on

Table 2: Summary of the Unfolded Matrices.

$\mathbf{M}_{t_k, \tau_m} =$	$\begin{matrix} (a, \rho) \\ \boxed{M} \\ \tau \end{matrix}$	$\begin{matrix} \text{Feature-channel matrix for each time window } t_k \text{ and delay pair } \tau_m \\ \text{(middle matrices in Figure 7)} \end{matrix}$
$\mathbf{M}_{b_l, t_k} =$	$\begin{matrix} \tau \\ \boxed{M} \\ \tau \end{matrix}$	$\begin{matrix} \text{Delay-channel matrix for each feature } b_l \text{ and time window } t_k \\ \text{(bottom matrices in Figure 7)} \end{matrix}$
$\mathbf{M}_{b_l, \tau_m} =$	$\begin{matrix} \tau \\ \boxed{M} \\ \tau \end{matrix}$	$\begin{matrix} \text{Time-channel matrix for each feature } b_l \text{ and delay pair } \tau_m \end{matrix}$
$\mathbf{M}_{C_n, t_k} =$	$\begin{matrix} (a, \rho) \\ \boxed{M} \\ \tau \end{matrix}$	$\begin{matrix} \text{Feature-delay pair matrix for each channel } C_n \text{ and time window } t_k \\ \text{(top matrices in Figure 7)} \end{matrix}$
$\mathbf{M}_{C_n, \tau_m} =$	$\begin{matrix} (a, \rho) \\ \boxed{M} \\ \tau \end{matrix}$	$\begin{matrix} \text{Time-feature matrix for each channel } C_n \text{ and delay pair } \tau_m \end{matrix}$
$\mathbf{M}_{C_n, b_l} =$	$\begin{matrix} \tau \\ \boxed{M} \\ \tau \end{matrix}$	$\begin{matrix} \text{Time-delay matrix for each channel } C_n \text{ and feature } b_l \end{matrix}$

all singular values obtained by SVD on combined unfolded matrices (see, e.g., Lathauwer et al., 2000, for details). In THOSVD only the first singular values on each of the red-framed matrices in Figure 7 are used. Table 2 summarizes the six unfolded matrices.

2.6 Delay Differential Analysis of ECoG Data. Delay differential analysis using the subset of model-delay pairs selected by the GA was performed with a significantly smaller 250 ms windows with 125 ms overlap. Thus, the GA was used to spiral search for the best-performing model-delay pairs, and the DDE was then run much more densely over the entire data set, increasing the temporal resolution and decreasing concerns associated with nonstationarity.

3 Results

3.1 Structure Selection. Although the delay differential and THOSVD analysis focused on two subjects, the GA was initially run on recordings from 13 subjects for each clinically defined seizure. The 13 subjects had 155 seizures and 730 ECoG channels, of which a subset of randomly selected

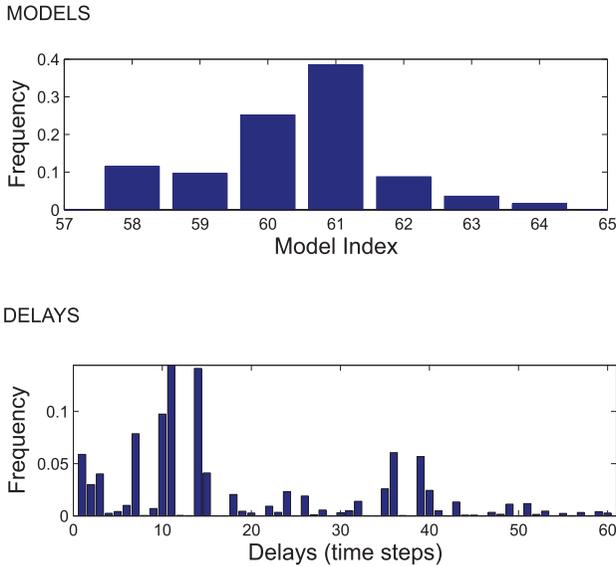


Figure 8: GA outputs based on ECoG from 13 subjects. Histograms of model (top) and delay (bottom) selection, using all channels for 30 minutes before and after seizure activity.

clinically defined seizure onset and nononset channels were run through the GA. GA analysis of 1 hour data segments, including 30 minutes prior and 30 minutes after the clinically defined onset allowed for a gross characterization of ECoG activity. The model and delay plots (see Figure 8) indicate that there is a subset of models and delays that appears to be selected most frequently when analyzing ECoG activity from the population with epilepsy. Based on the delay selection (see Figure 8, bottom), the following delay pairs were selected for use in the DDA: (1, 2), (1, 3), (7, 10), (11, 14), (15, 18), (24, 26), (36, 39), (35, 40). These delay pairs will be referred to as delay pair 1, 2, . . . , 8 for the remainder of the analysis. The model selection (see Figure 8, top) shows that the model indices 58 to 64, representing the models

$$\text{model 58: } \dot{x} = x_{\tau_1} + x_{\tau_2} + x_{\tau_1}^2$$

$$\text{model 59: } \dot{x} = x_{\tau_1} + x_{\tau_2} + x_{\tau_1}x_{\tau_2}$$

$$\text{model 60: } \dot{x} = x_{\tau_1} + x_{\tau_2} + x_{\tau_1}^3$$

$$\text{model 61: } \dot{x} = x_{\tau_1} + x_{\tau_2} + x_{\tau_1}^2x_{\tau_2},$$

$$\text{model 62: } \dot{x} = x_{\tau_1} + x_{\tau_2} + x_{\tau_1}^4$$

$$\begin{aligned}
 \text{model 63: } \dot{x} &= x_{\tau_1} + x_{\tau_2} + x_{\tau_1}^3 x_{\tau_2} \\
 \text{model 64: } \dot{x} &= x_{\tau_1} + x_{\tau_2} + x_{\tau_1}^2 x_{\tau_2}^2
 \end{aligned}
 \tag{3.1}$$

were selected most frequently. Figure 8 is focused on the model indices 58 to 64 because it was rare that any models outside this range were selected by the GA. Any of the identified models could be selected for further analysis as each provides different types of temporal-scale and higher-order statistical information about the ECoG signal.

The top panel of Figure 9 shows an example of the model selection for patient 1 for one electrode for one seizure. Models 58 and 62 show differences before and after seizure onset (red line). Based on this observation and the model histogram in Figure 8, model 62, $\dot{x} = x_{\tau_1} + x_{\tau_2} + x_{\tau_1}^4$, was selected under the assumption that the inclusion of the higher-order nonlinear term would increase the retention of higher-order statistical information.

3.2 Delay Differential Analysis and THOSVD. After the GA structure selection, DDA was performed with the model and delay pairs that were consistently selected during common ECoG events, including artifacts, stimulations, interictal activity, electroclinical seizures, and subclinical seizures. After obtaining the DDA features a_1, \dots, ρ for each of the events, THOSVD was applied, allowing for the comparison of the different tensor unfoldings for each of the selected events. The figures shown here for the DDA analysis will be limited to the first DDE feature a_1 due to space considerations. However, all of the DDA features were computed and appropriately utilized for THOSVD computation as described previously.

The different tensor unfoldings (see Table 2) dichotomize into two separate but related subgroups. The first group, global-THOSVD, performs SVD on unfolded matrices that have the channel location C_n as one of the dimensions of matrix \mathbf{M} . Performing SVD on channel-x combinations, where x corresponds to features, delays, or time, results in a most prominent singular value that is completely compressed across the spatial dimension. Thus, the most prominent singular value is considered a spatially global characteristic. Dichotomously, the local-THOSVD retains all of the spatial information as the tensor is unfolded such that the channels C_n are not included in the dimensions of matrix \mathbf{M} . Thus, the spatial information is retained and the matrix is instead compressed across combinations of time, delays, and features. Importantly, these subgroups of THOSVD analysis are related, as both are built on the same DDA-feature tensor but compressed across different dimensions.

3.2.1 THOSVD Analysis for a Seizure. The singular values of the unfolded matrices summarized in Table 2 were computed, and Figure 10 displays examples of $SV(\mathbf{M}_{b_l, t_k})$ for clinical and electrographic seizures respectively.

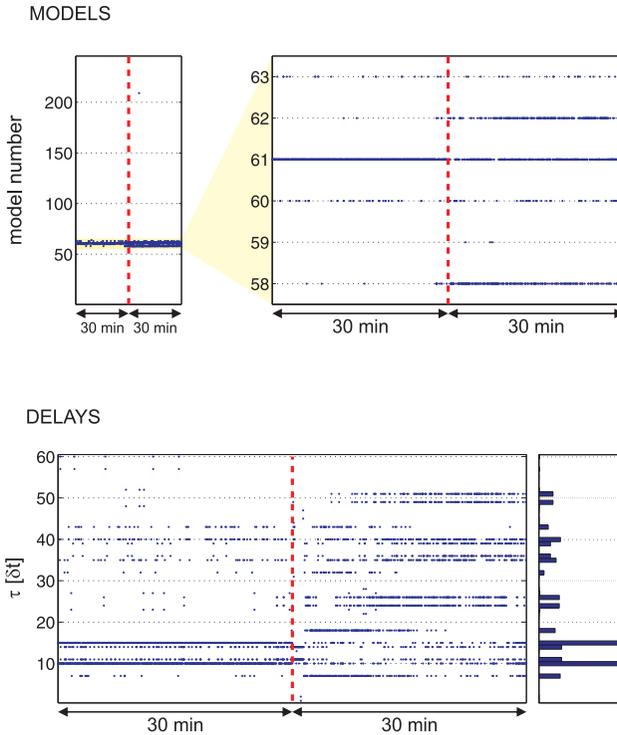


Figure 9: Example of model and delay selection outputs for one electrode for one seizure in patient 1 showing evidence for a bifurcation in the cortical state around the seizure. (Top) Model plots displaying model were chosen with respect to time. On the left, model selection showing robust selection in the model index band comprising indices 58 to 63. On the right, a zoomed-in version of the model plot displaying the model selection during the interictal, pre-ictal, ictal, and post-ictal periods. Model 61 was chosen in all periods, but models 58 and 62 were chosen only on the post-ictal period. (Bottom) Delay selection from GA output. Each point pair represents a pair of delays selected for a given time window. Pre-ictal delay selection includes 10, 11, 14, and 15 δt . The ictal state most commonly selects for delays of 11 and 14 δt . Finally, the post-ictal states selects for delays of 7 and 18 δt . The plot on the right shows a histogram of the selected delays for the entire hour recording.

The seizures are marked by orange outlines in the scatter plots. Approximately 100 channels are displayed in each plot. To see the channels involved in the seizure and, more important, the onset channels, these plots are converted into heat maps where the value of the singular value is then shown as color (see Figure 10). The seizures in the heat maps are marked with pink

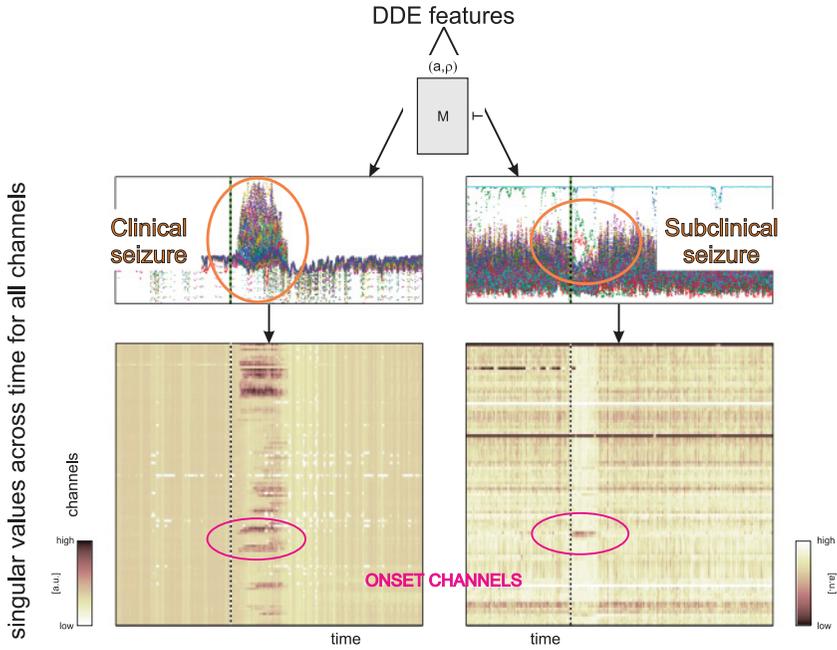


Figure 10: Sketch of the THOSVD analysis pipeline for the identification of seizure onset and termination. The DDA feature tensor in equation 2.11 is unfolded into a matrix M for each channel. All of the singular value outputs for each subject’s ECoG channels are plotted together in the middle row scatter plots, and the seizures are outlined in orange. Many channels have strong signals following the onset of the clinical seizure but have reduced activity for the subclinical seizure. The outputs are spatially represented in all the channels in the bottom row heat maps, with colors ranging from white to brown where the seizures are marked with pink circles.

circles. Note that the clinical seizure was more widespread and that the subclinical seizure was more localized and occupied by a decrease in activity in the other channels.

3.2.2 *Topographical Brain Reconstruction.* In order to provide a sense of how different ECoG events manifest spatially across the cortex, the topographical cortical reconstructions of patients 1 and 2 are provided in Figure 11. The reconstructions were performed through the use of preoperative MRI and postoperative high-resolution CT. First, three-dimensional renderings of each patient’s cortical surface were generated utilizing the high-resolution MRI. Postoperative CT scan (e.g., after ECoG grid and electrode implantation) was then used to localize the electrodes using a previously

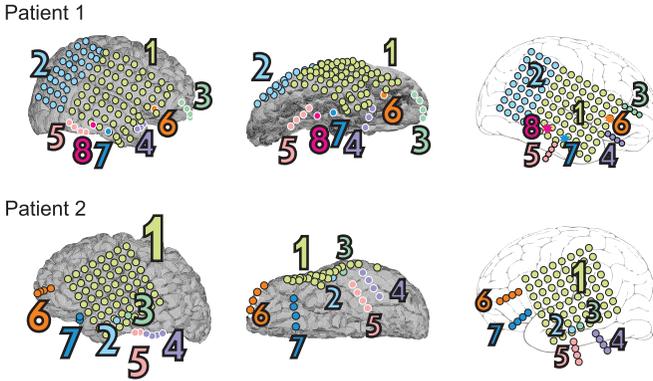


Figure 11: Approximate reconstruction of the cortical topography of patients 1 and 2 showing the ECoG locations superimposed. The cortical topographical reconstructions are displayed with lateral (left), inferior (middle), and schematic lateral (right) views. The grid area colors in the reconstruction are identical to the colors in the simplified graphical representation shown at the right for both patients. Each group of electrodes is indicated by a colored number and a corresponding array of circles. For patient 1, electrode arrays 1 and 2 are grid electrodes, 7 and 8 are depth electrodes, and all others are strip electrodes. For patient 2, array 1 is a grid, 2 and 3 are depth electrodes, and all others are strip electrodes.

validated energy optimization algorithm (Dykstra et al., 2012). The remaining analysis focuses on patients 1 and 2, and spatial localization of events across the cortical surface should refer to the reconstructions in Figure 11.

3.2.3 Ictal Events. Figure 12 displays the DDE feature a_1 for an electroclinical seizure (patient 1) and a subclinical seizure (patient 2). The time course of a_1 in the electroclinical seizure clearly shows the beginning, evolution, and termination of the seizure in both the scatter plot (top panel) and the heat maps (bottom panel). The clinical seizure quickly evolves from a localized event in region 2 to a generalized event prior to terminating. Interestingly, the seizure appears to terminate simultaneously across the entire grid, indicating global activity at seizure termination. In contrast, the electrographic seizure shows highly localized activity in region 2 (temporal) with additional activity in regions 1, 3, and 7. Region 2 corresponds to a depth electrode, and the increased activity in regions 1, 3, and 7 corresponds to the portions of these strip and grid electrodes that surround the depth electrode. As such, it is difficult to ascertain if there is true ictal activity in these electrodes or if this is simply electrical spreading from an epileptogenic focus in region 2.

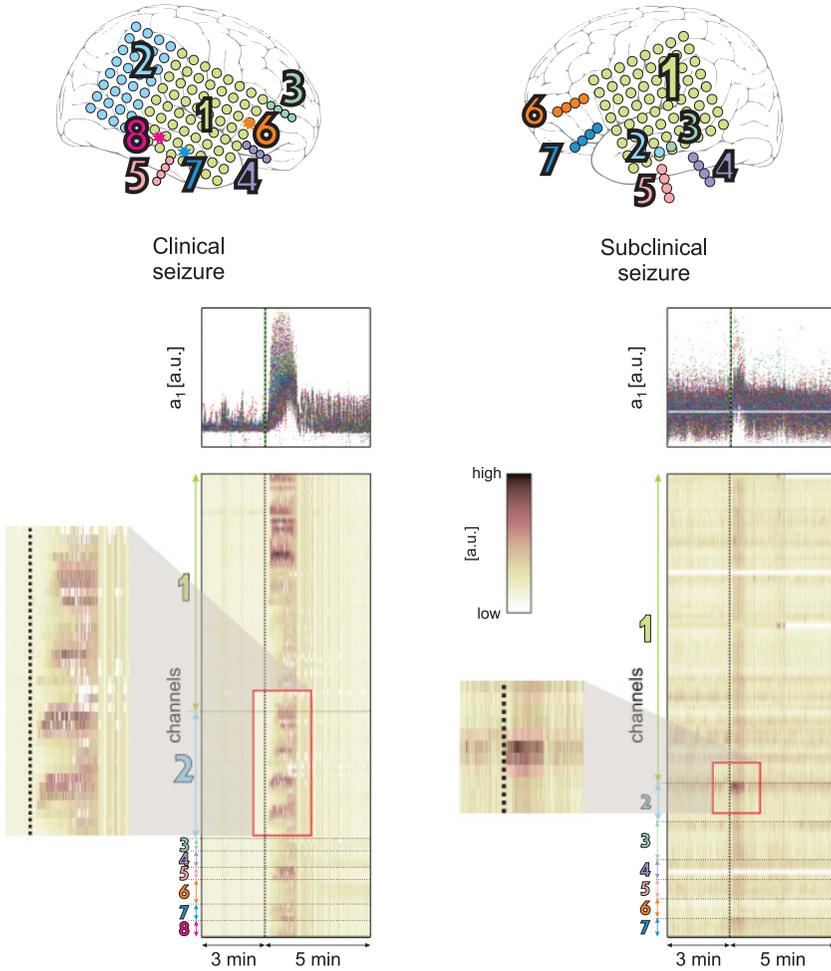


Figure 12: DDE feature a_1 for an electroclinical seizure (left) and subclinical seizure (right). The patient-specific electrode maps are displayed in the top panel. Scatter plots (middle panel) and heat maps (lower panel) corresponding to the two events provide insight to the onset, evolution, and termination of the events.

Figure 13 displays the singular values with retention of spatial information, $SV(\mathbf{M}_{t_k, \tau_m})$, $SV(\mathbf{M}_{b_l, t_k})$, and $SV(\mathbf{M}_{b_l, \tau_m})$, for clinical and subclinical seizures. The singular values for the clinical seizure clearly display the onset, evolution, and termination of the ictal activity in a spatially localized manner. Notably, the singular values provide a cleaner representation of the ictal activity than the DDE features a_1 (see Figure 12) with the best

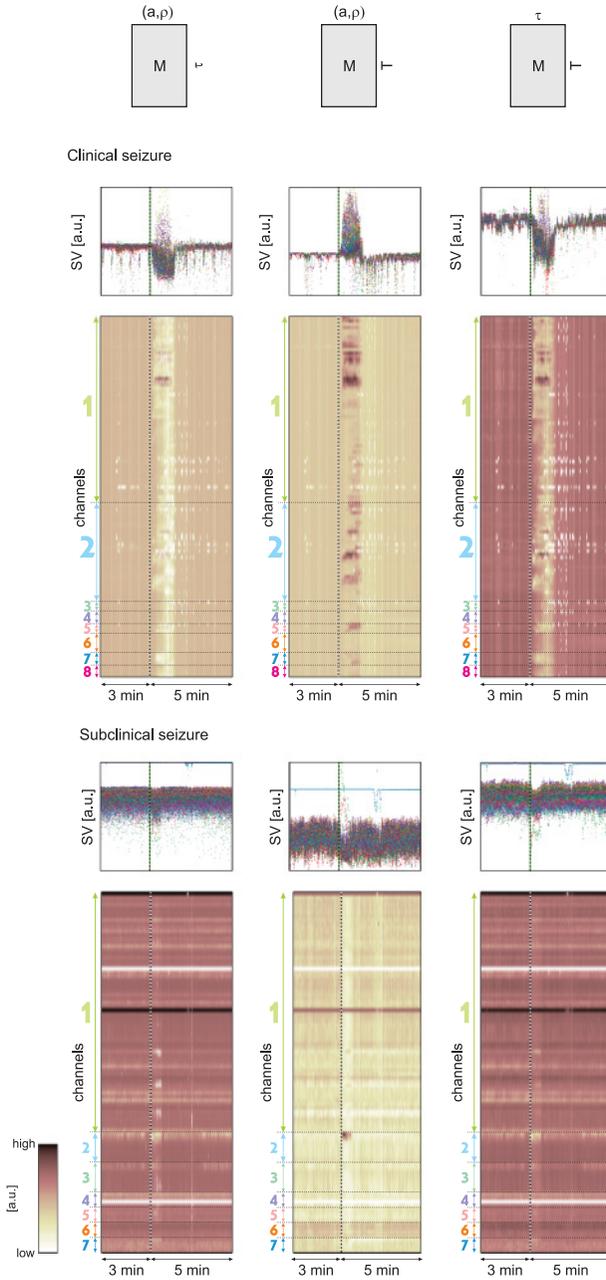


Figure 13: Singular values with retention of spatial information, from left to right; $SV(\mathbf{M}_{k, \tau_m})$, $SV(\mathbf{M}_{b_j, t_k})$, and $SV(\mathbf{M}_{b_j, \tau_m})$, of a clinical seizure (top panel) and subclinical seizure (bottom panel) from patients 1 and 2, respectively.

representation in $SV(\mathbf{M}_{b_l, t_k})$. Importantly, the localized singular values also robustly describe the ictal activity for the subclinical seizure. While there is activity throughout the grid prior to the electrographic seizure, there is highly localized activity in region 2 during the ictal event with very limited activity elsewhere. Thus, the onset, evolution, and termination are again well displayed, and a robust view of the epileptogenic focus is obtained. With further inspection, it is clear that $SV(\mathbf{M}_{t_k, \tau_m})$ responds quite poorly to the electrographic seizure, likely because the activity is highly localized and largely absent from the DDE features, a_2, a_3 , and ρ (not shown). Both the $SV(\mathbf{M}_{b_l, t_k})$ and $SV(\mathbf{M}_{b_l, \tau_m})$ display a better response because the electrographic seizure is plotted solely for a_1 and the singular values do not utilize the information in a_2, a_3 , or ρ .

The singular values compressed across the spatial dimension are shown in Figure 14. For the clinical seizure, the global singular values further enforce that this is a global, large-scale event that dominates with both positive and negative deflections for the various sets of delay pairs. Each of the delay pairs deflects away from its baseline and returns to its baseline at roughly the same time points. This indicates a global onset and offset for the seizure activity that is robustly identified across the spatial scales. Similar to the electroclinical seizure, the electrographic seizure is best represented by $SV(\mathbf{M}_{C_n, \tau_m})$, which shows a global state change and negative deflection. $SV(\mathbf{M}_{C_n, \tau_m})$ was plotted only for the first delay pair, and when comparing $SV(\mathbf{M}_{C_n, \tau_m})$ to $SV(\mathbf{M}_{C_n, \tau_m})$ and $SV(\mathbf{M}_{C_n, b_l})$, it is clear that the robust response to the seizure is dependent on the selection of the proper delay pair, which is notable in that it contrasts significantly with the clinical seizure. Thus, the singular value computation across time and features provides a highly descriptive characterization of the electrographic activity while also filtering out noise processes that are not related to the seizure.

3.2.4 Additional Events. During any biophysical recording, it is likely that a wide array of artifacts will be recorded. As such, it is necessary that any data processing technique that aims to classify or characterize a system must be able to robustly discriminate between artifacts, stimulations, or general interictal periods and events of interest. Figure 15 displays the DDE feature a_1 for stimulation, artifact, and interictal events in ECoG recordings obtained from patient 1. The DDE feature a_1 qualitatively shows the artifact to be a global event that is easily differentiated from the electroclinical, electrographic, stimulation, and interictal events. The feature's response to exogenous stimulation shows a global negative deflection with a unique onset, time course, and termination when comparing with the other ECoG events.

The length of traditional ECoG recordings obviates the ability to distinguish interictal activity from other events of interest. Comparing the interictal feature a_1 time course to the other ECoG event recordings leaves the impression that there is no immediately discernible structure. The lack of

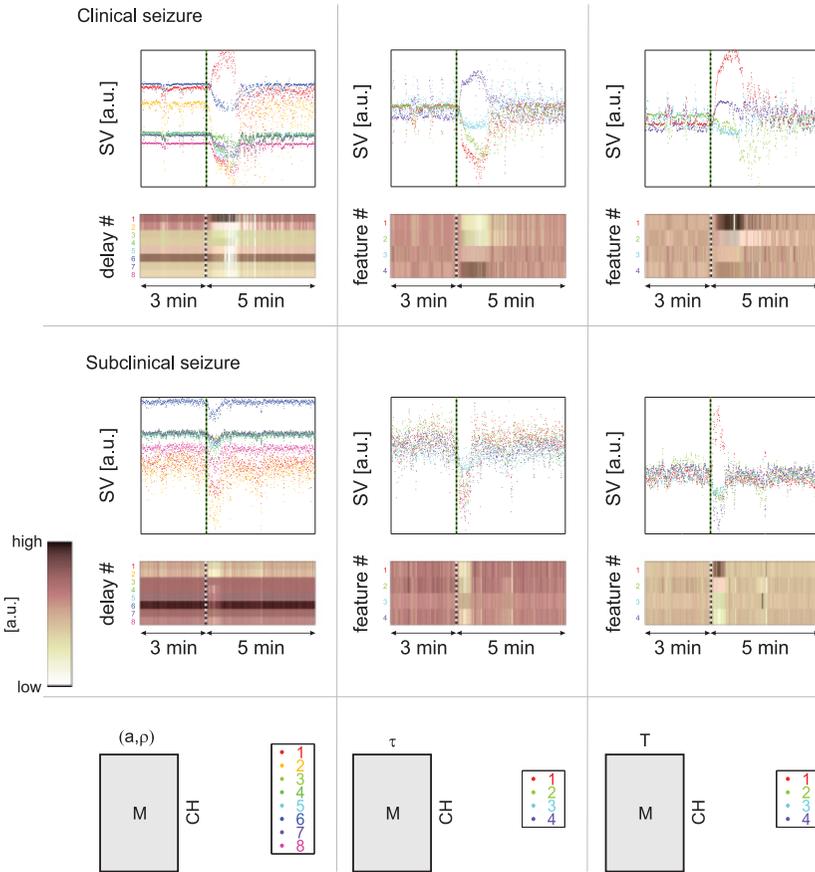


Figure 14: Singular values with compression of spatial information, (from left to right) $SV(\mathbf{M}_{C_n, b_l})$, $SV(\mathbf{M}_{C_n, \tau_m})$, and $SV(\mathbf{M}_{C_n, t_k})$, of a clinical seizure (top panel) and subclinical event (middle panel) from patients 1 and 2 respectively. Scatter plots and heat maps are displayed for each seizure example. Delay pair, feature indices, and matrix icons are shown in the lower panel.

apparent structure and overall irregularity of the signal makes it readily differentiable from the aforementioned ECoG events. However, the lack of structure given this specific DDA model-delay pair should not lead to the conclusion that there is no structure or dynamical information within the interictal recording. Such analysis will be the subject of future work, but it is important to remember that all ECoG signals, regardless of the event taking place, contain dynamical information.

Singular values compressed across the spatial dimension and with retention of the spatial dimension for artifacts, exogenous stimulations, and

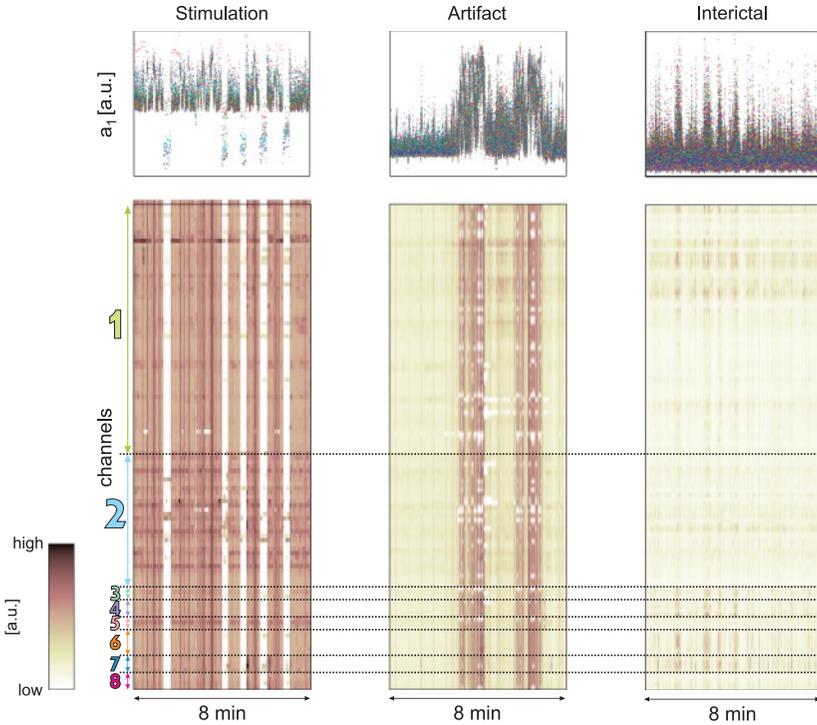


Figure 15: DDE feature a_1 for (from left to right) stimulation, artifact, and interictal events. Scatter plots (top panel) and heat maps (middle panel) for each of the two events provide a graphical representation of the onset, evolution, and termination of the events recorded from patient 1.

interictal periods are shown in Figure 16. The global singular values describe the artifact as a global process with dependence on feature selection (e.g., different features display the artifact with positive and negative deflections). This is to be expected as different dynamical and temporal scale information is contained within each feature. The local SVs show the artifact to be quite noisy, with a global downward deflection in its first and second phases. The singular value outputs for exogenous electrical stimuli indicate that global negative deflections are best visualized using the global singular value $SV(\mathbf{M}_{C_n, \tau_m})$. The scatter plot and heat map representations, $SV(\mathbf{M}_{b_l, t_k})$, of the exogenous electrical stimuli show widespread events with negative deflections in all channels when compressing across features and time. Similar to the DDE representation, the global and local SVs for the interictal period appear quite noisy and display no large-scale events.

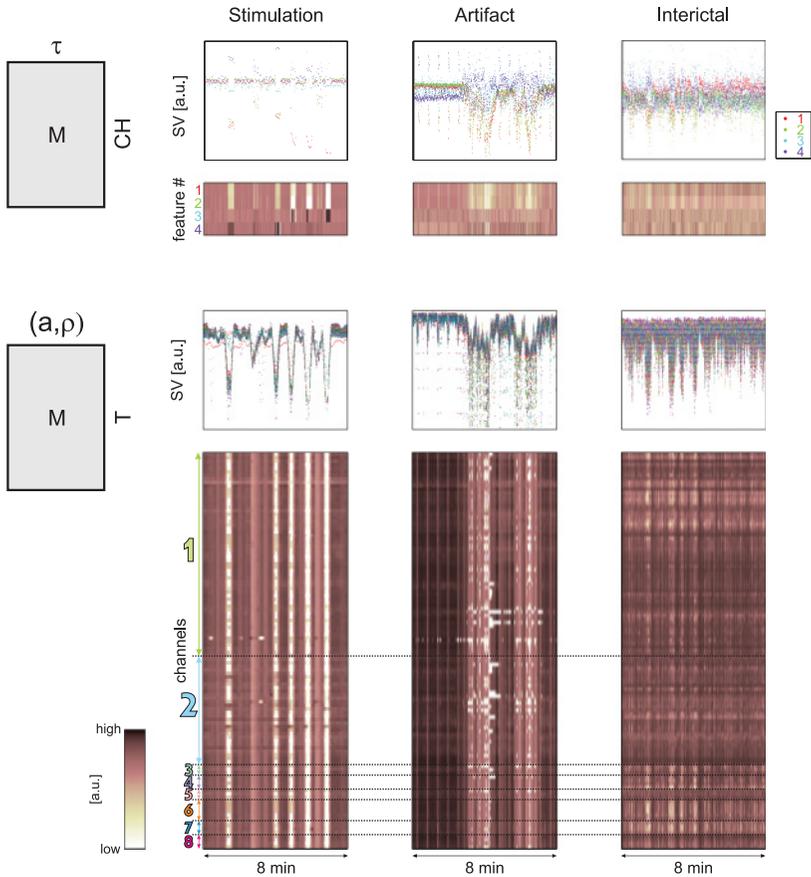


Figure 16: Singular values $SV(\mathbf{M}_{b_l, t_k})$ (top panel) and $SV(\mathbf{M}_{C_n, \tau_1})$ (bottom panel) for the ECoG events corresponding to exogenous electrical stimulation, artifacts, and interictal periods from patient 1.

4 Future Directions

4.1 DDA and Clinical Findings. The ECoG data from patient 1 were initially examined by human experts in a clinical setting for the identification of epileptogenic zones and subsequent operative planning. This patient had five seizures identified in the clinical report with varying onset channels for the individual ictal events: seizures 1 and 2 were marked to start in anterior temporal depth channels; seizures 3 and 5 were marked to start in posterior temporal grid channels; for seizure 4, a “broad onset–sentinal spike” was noted (see Table 3 and Figure 17). DDA analysis found two additional seizures: one after a stimulation (seizures 0 in Figure 17) and one

Table 3: Clinical Seizures of Patient 1 and Onset locations as Determined by a Human Clinical Expert.

Seizure Number	Onset Channels
1	ATD1,2,3
2	ATD1,2,3
3	PsGr19, 26, 27, 34, 43, 44
4	Broad onset-sential spike
5	PsGr19, 26, 27, 34, 43, 44

Note: Two seizures start in the posterior temporal lobe, two start in the anterior temporal lobe, and for one, no onset locations can be determined.

that was not fully developed (seizures 6 in Figure 17). These seizures were identified by characterizing the known seizures with DDA-THOSVD, with subsequent implementation of computer recognition of similar events. As can be seen in Figure 17, all seizures have the same onset channels when characterized with the DDA-THOSVD technique. Movies on the temporal evolution of these seizures are provided in the supplemental material (<http://snl.salk.edu/~claudia/EpilepsyMovies>). Onset times slightly differed from the marked onset times (see the vertical lines in Figure 17). Importantly, both the additional seizures and DDA-THOSVD onset times were confirmed retrospectively by a human expert.

Patient 2 had 74 seizures of which 2 were clinical. Figure 18 shows the two clinical seizures (36 and 45) and five subclinical seizures. From these figures, the consistency across all seizures is readily discernable.

These examples and a complete analysis of all 15 patients needs a detailed report, which is in preparation.

4.2 Comparison with Spectral Features. As mentioned in section 5, DDA, spectral analysis, and higher-order statistics are interconnected (Lainscsek & Sejnowski, 2015). In Figure 17 we showed the similarity between the clinical seizures of patient 1 despite the differences in the clinical report as reported in the previous section. For a more complete analysis, spectral features should not be neglected. Schiff et al. (2000) analyzed time-frequency spectrograms of epileptic seizures and found that frequency chirps are very sensitive detectors of seizures (83%) and highly specific as markers with no false-positive detections. In Figure 19 spectrograms of patient 1 are shown for the onset channel. We see the brain chirps in all seizures. As introduced in Lainscsek and Sejnowski (2015), we also computed time-frequency bispectrograms as shown in Figure 20 of the same seizures to assess the relevance of nonlinearities in the data. We noticed bispectral brain chirps indicating the need for nonlinear analysis, such as DDA. Additionally, (bi-)spectrograms are high-dimensional across time, while DDA yields

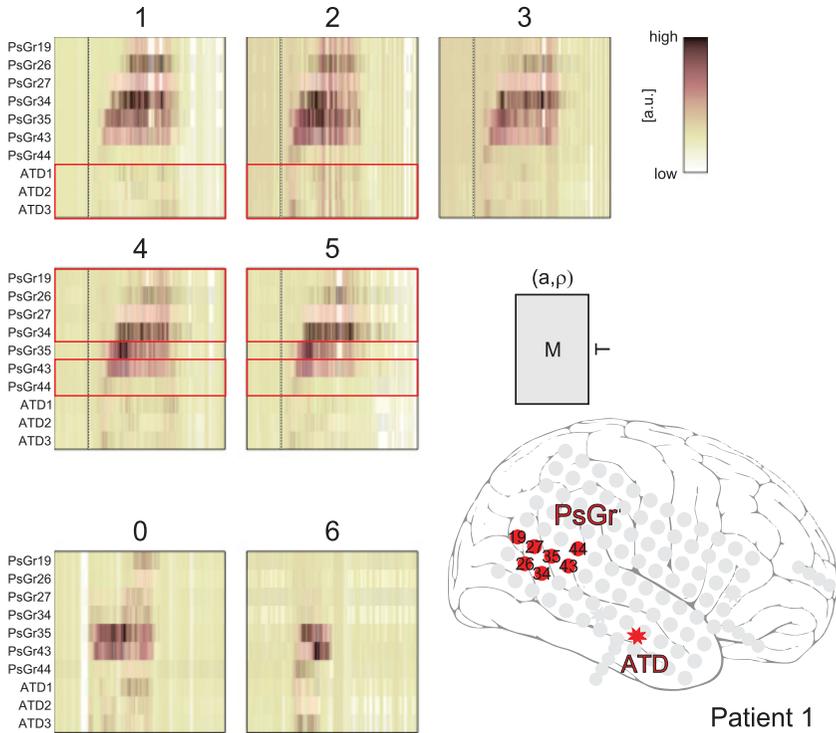


Figure 17: Clinical seizures from patient 1: seizures 1 to 5 were described in the clinical report with the onset locations summarized in Table 3 as determined by a human expert. Three minutes of data are shown—0.5 minute before the onset time determined by a human expert (vertical lines) and 2.5 minutes after the onset time. The onset locations from the report are highlighted with red boxes. Two additional seizures (seizures 0 and 6) were found by DDA. Seizure 0 started right after a stimulation (light vertical area) and one was not fully developed. DDA analysis shows very similar temporal evolution of all seizures.

low-dimensional features that can be easily mapped into a one-dimensional index as outlined in this article.

A detailed analysis of these findings is subject to ongoing research and will be published later.

5 Discussion

Perhaps the most surprising result of our analysis was that single DDEs provided good fits to the ECoG waveforms of all our subjects. This suggests that the terms in these DDEs reflect basic features of cortical dynamics that

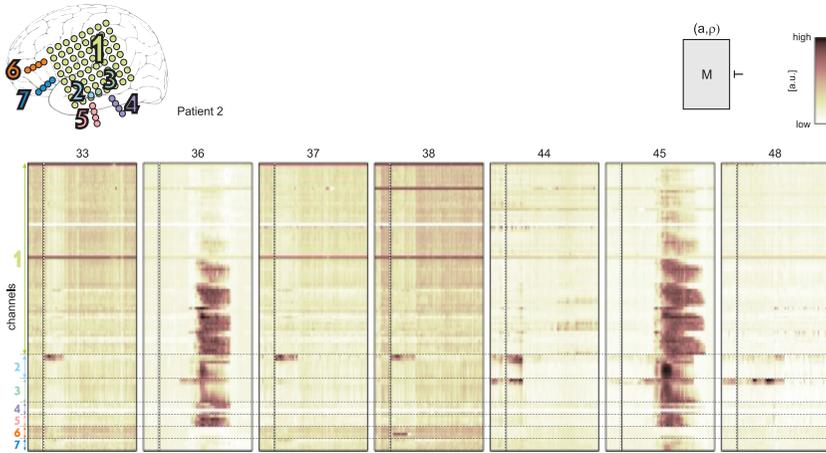


Figure 18: Clinical (36 and 45) and subclinical seizures from patient 2. Three minutes of data are shown—half a minute before the onset time determined by a human expert (vertical lines) and 2.5 minutes after the onset time. DDA analysis shows the similarity of all seizures.

are universal to the population with epilepsy in this study. It must be noted that the patients with epilepsy in this study represented multiple underlying pathologies including mesial temporal sclerosis, cortical dysplasia, and traumatic epilepsy syndromes. As such, the analysis described enables the construction of a limited feature space—combinations of eight delays and six DDE models—that can be used to robustly describe pre-ictal, ictal, and post-ictal activity irrespective of the underlying etiology or spatial scale of a patient’s epileptic activity.

In addition, we have shown that each type of ECoG event has its own characteristic DDE pattern across the two subjects whose ECoG signals were analyzed in depth. Furthermore, the selection of features and subsequent THOSVD compression allows for specific properties of the ECoG events to be used for characterization and identification. Indeed, all of the events are qualitatively different from all of the others and can be identified by simple pattern recognition, which suggests that the analysis can be automated. This will be the subject of future work in a larger patient cohort allowing for a more comprehensive statistical analysis. This article is meant to describe the DDA-THOSVD technique and its implications in ECoG analysis.

The ability to robustly describe interictal activity, irrespective of epilepsy syndrome, with a limited feature space indicates that the dynamical characteristics of epilepsy syndromes at the scale of ECoG recordings are quite similar. Explicitly, within the interictal period, there are specific nonlinear

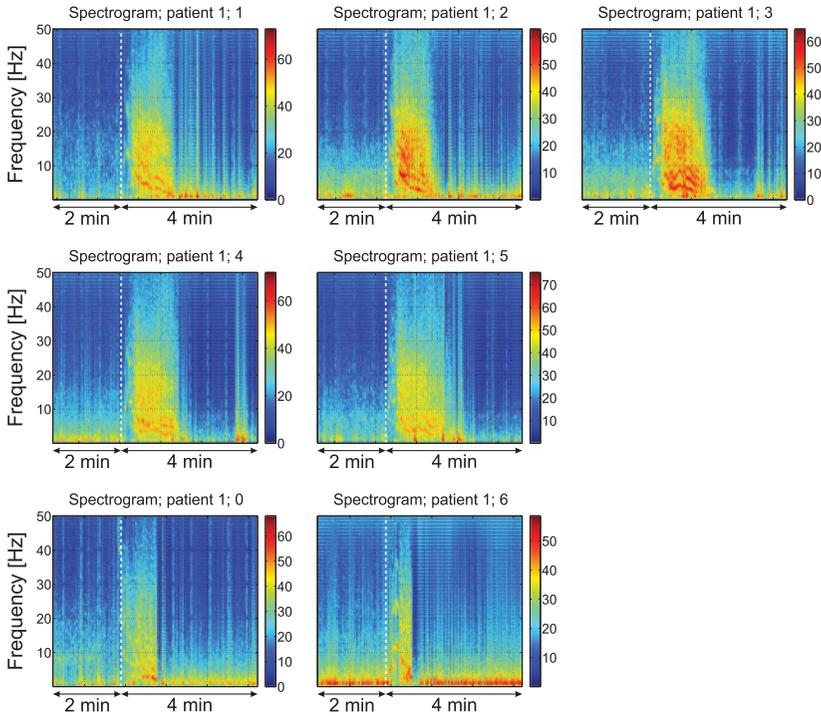


Figure 19: Spectrograms of the clinical seizures of patient 1.

DDA features that robustly describe ECoG waveforms invariant of subject. This does not mean that nonpathological cortical activity is inherently simple or can be described by the same feature set. Rather, the DDA models and delays were selected for their ability to adequately describe the interictal period and to differentiate between the pre-ictal, ictal, and post-ictal periods using a genetic algorithm based search method and ECoG data that had been previously annotated for seizure onset time and channels. However, this analysis technique is readily expandable to other tasks, such as sleep stage classification or neurological disease classification and gradation, at varying spatial scales within the cortical recording spectrum (Lainscsek et al., 2014; Lainscsek, Hernandez et al., 2013).

Results from previous EEG analyses with DDA (Lainscsek et al., 2014; Lainscsek, Hernandez et al., 2013) have shown that specific model structures, such as combinations of two linear terms and one nonlinear term, robustly classify neurological events and pathologies. Analogously, similar models were selected for this data set (see Figure 8 and equation 3.1). In Lainscsek and Sejnowski (2015) connections between DDA, spectral analysis, and higher-order statistics were presented. A linear DDE can be used to

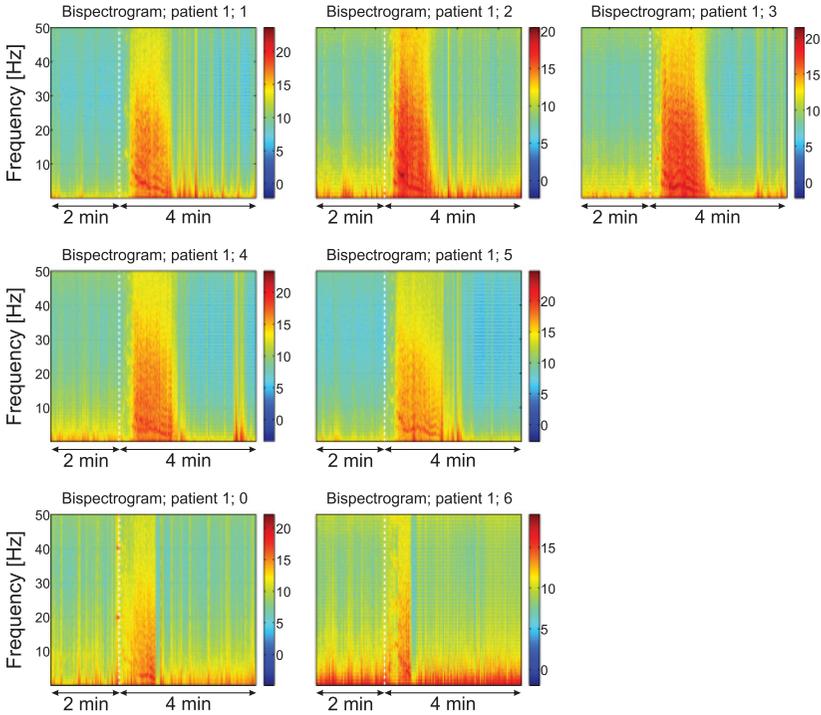


Figure 20: Bispectrograms of the clinical seizures of patient 1.

do spectral analysis, whereas a DDE with a quadratic term corresponds to a bispectrum. For DDEs with linear and nonlinear terms, spectral features and higher-order terms are interconnected. The simple connections between frequencies, frequency, couplings, and DDE features are lost. Such a trade-off between good spectral certainty and good detection rate can be interpreted as an uncertainty principle.

The goal of this analysis was not only to identify different types of ECoG activity but also to begin to explore the underlying dynamics involved in these activities in order to further understand pathological brain states, specifically with respect to epilepsy. This was accomplished by dissecting the DDA outputs using THOSVD in order to emphasize the different types of dynamical information present in the DDA feature space. For example, the inclusion of features related to a specific delay pair will isolate specific timescales, and only events with the timescales of interest will show a feature response. The selection of specific temporal scales does not necessarily map to a specific dominant frequency in the ECoG signal but to a higher-order statistical frequency-phase combination consisting of multiple frequencies and phases that are coupled to produce ECoG phenomena.

Furthermore, the inclusion of all of the coefficients will result in the retention of information related to all of the frequencies in the ECoG signal (see equation 2.8). Conversely, the choice to retain only the nonlinear coefficient will emphasize the higher-order correlations present in the signal. A robust dynamical characterization of ECoG requires the combination of spatial, temporal (frequency), and higher-order statistical information in order to identify specific dynamical characteristics found in the different events. As with many machine learning paradigms, it is difficult to relate the DDA features to components of the analyzed signal that can be discerned with the naked eye or readily understood within the traditional Fourier frequency analysis framework. This caveat arises due to the retention of information that is readily ignored in traditional frequency analysis or invisible in the time-amplitude space.

The singular value analysis is a change of basis that isolates specific types of DDA information to allow a more compact understanding of the underlying dynamical system. The global singular values are a window into the global dynamics of the cortical state and events of interest. If an ECoG event has global dynamics during onset, evolution, or termination, then these dynamics are readily discernible with the global THOSVD analysis. In addition, the global dynamic representation by a specific feature, delay pair, or segment of time provides a specific tool for identifying the dynamics and timescales involved in the generation and termination of cortical events. Although the structure selection and delay differential analysis were subject invariant, this portion of the analysis is subject and event specific as it clearly separates the different events based on the information variables that best characterize the underlying dynamics.

The local singular values are a lens that can discern localized dynamics and smaller events. For example, the scatter plots in Figure 13 pick out a small event that may have occurred after a clinically defined seizure; however, it is difficult to see the details. A much better representation is displayed in the heat maps in Figure 13, where the seizure is clearly localized to a subset of electrodes. The seizure activity is picked up most strongly when combining both the temporal and feature information variables and choosing a specific delay pair. Thus, each electrographic seizure has dynamics that are best characterized by a single delay pair. Interestingly, the nonclinical seizure also has global dynamics (see Figure 14), even though the actual ictal activity was highly localized. This finding makes it likely that global processes are involved in its generation and termination, a finding consistent with a growing body of literature focusing on the network activity inherent in ictal generation. This warrants further study as it may indicate that subclinical seizures are the result of global processes influencing a localized section of parenchyma.

Local THOSVD analysis can identify spatially localized events, such as seizure onset, seizure evolution, interictal spikes, and sleep spindles. The local singular values are sensitive to the temporal dynamics and higher-order

statistical information contained in each channel prior to, during, and after each event. The high temporal resolution that is possible in both global and local dynamics is a significant strength of the THOSVD method in delineating state transitions, specifically as they relate to the interictal and ictal periods. The high temporal resolution of highly localized ECoG DDA features should provide a better dynamical understanding of the onset and termination of ECoG events and can be compared and integrated with emerging microscale studies of single neuron activity during seizure initiation, propagation, and termination (Truccolo et al., 2011, 2014; Smith et al., 2016). Furthermore, the ability to dissect the dynamics of seizures will lead to improved localization and characterization of individual seizures both within and across subjects and epilepsy syndromes. An in-depth clinical analysis of these subjects will be presented elsewhere.

6 Conclusion

We have described a new method for the characterization of ECoG states and cortical events based on methods from dynamical systems theory. The data-reduction techniques are compatible with big data and make dynamical systems methods accessible to the clinic. We have demonstrated the ability to qualitatively discriminate between electroclinical events and sub-clinical seizures based on their dynamical and spatial properties. The dynamical and spatial information within the signal can be used to further understand the mechanisms underlying the dynamics of seizure initiation and termination.

Acknowledgments

We thank Aaron Sampson for helpful discussions. This work was supported by the Howard Hughes Medical Institute, NIH (grant NS040522) Office of Naval Research (N000141210299), the Swartz Foundation, and NIH (grant NS062092) to S.S.C.

References

- Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P., & Elger, C. E. (2001). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E*, *64*, 1–8.
- Broomhead, D., & King, G. (1986). Extracting qualitative dynamics from experimental data. *Physica D*, *20*, 217.
- Buzsaki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and current EEG, ECoG, LFP, and spikes. *Nature Reviews Neuroscience*, *13*, 407–420.

- Cover, T. (1965). Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers, EC-14*(3), 326–334.
- da Silva, F. L., Blanes, W., Kalitzin, S. N., Parra, J., Suffczynski, P., & Velis, D. N. (2003). Epilepsies as dynamical diseases of brain systems: Basic models of the transition between normal and epileptic activity. *Epilepsia, 44*, 72–83.
- Dominguez, L. G., Wennberg, R. A., Gaetz, W., Cheyne, D., Snead III, O. C., & Velazquez, J. L. P. (2005). Enhanced synchrony in epileptiform activity? Local versus distant phase synchronization in generalized seizures. *Journal of Neuroscience, 25*(35), 8077–8084.
- Dykstra, A. R., Chan, A. M., Quinn, B. T., Zepeda, R., Keller, C. J., Cormier, J., ... Cash, S. S. (2012). Individualized localization and cortical surface-based registration of intracranial electrodes. *NeuroImage, 59*, 3563–3570.
- Elger, C. E., & Lehnertz, K. (1998). Seizure prediction by non-linear time series analysis of brain electrical activity. *European Journal of Neuroscience, 10*, 786–789.
- Goldberg, D. (1998). *Genetic algorithms in search, optimization and machine learning*. Reading, MA: Addison-Wesley.
- Hornero, R., Abasolo, D., Escudero, J., & Gomez, C. (2009). Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease. *Phil. Trans. R. Soc. A, 367*, 317–336.
- Iasemidis, L. D., & Sackellares, J. C. (1996). Chaos theory and epilepsy. *Neuroscientist, 2*, 118–126.
- Jiruska, P., de Curtis, M., Jefferys, J. G. R., Schevon, C. A., Schiff, S. J., & Schindler, K. (2012). Synchronization and desynchronization in epilepsy: Controversies and hypotheses. *Journal of Physiology, 591*, 787–797.
- Judd, K., & Mees, A. (1995). On selecting models for nonlinear time series. *Physica D, 82*, 426–444.
- Kadtke, J., & Kremliovsky, M. (1997). Estimating statistics for detecting determinism using global dynamical models. *Phys. Lett. A, 229*(2), 97.
- Kramer, M. A., Truccolob, W., Edena, U. T., Lepagea, K. Q., Hochberg, L. R., Eskandar, E. N., ... Cash, S. S. (2013). Human seizures self-terminate across spatial scales via a critical transition. *PNAS, 109*, 21116–21121.
- Kremliovsky, M., & Kadtke, J. (1997). Using delay differential equations as dynamical classifiers. In *Proceedings of the Conference on Applied Nonlinear Dynamics near the Millennium*. College Park, MD: American Institute of Physics.
- Lainscsek, C., Hernandez, M. E., Weyhenmeyer, J., Sejnowski, T. J., & Poizner, H. (2013). Non-linear dynamical analysis of EEG time series distinguishes patients with Parkinson's disease from healthy individuals. *Frontiers in Neurology, 4*(200), 1–8.
- Lainscsek, C., Messenger, V., Portman, A., Sejnowski, T. J., & Letellier, C. (2014). Automatic sleep scoring from a single electrode using delay differential equations. In J. Awrejcewicz (Ed.), *Applied non-linear dynamical systems* (pp. 371–382). New York: Springer.
- Lainscsek, C., Rowat, P., Schettino, L., Lee, D., Song, D., Letellier, C., & Poizner, H. (2012). Finger tapping movements of Parkinson's disease patients automatically rated using nonlinear delay differential equations. *Chaos, 22*, 013119.

- Lainscsek, C., & Sejnowski, T. (2013). Electrocardiogram classification using delay differential equations. *Chaos*, 23(2), 023132.
- Lainscsek, C., & Sejnowski, T. (2015). Delay differential analysis of time series. *Neural Computation*, 27(3), 594–614.
- Lainscsek, C., Weyhenmeyer, J., Hernandez, M., Poizner, H., & Sejnowski, T. (2013). Non-linear dynamical classification of short time series of the Rössler system in high noise regimes. *Frontiers in Neurology*, 4(182).
- Lathauwer, L. D., de Moor, B., & Vandewalle, J. (2000). A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.*, 21(4), 1253–1278.
- Lehnertz, K., & Elger, C. E. (1998). Can epileptic seizures be predicted? Evidence from nonlinear time series analysis of brain electrical activity. *Physical Review Letters*, 80, 5019–5022.
- Lehnertz, K., Mormann, F., Kreuz, T., Andrzejak, R. G., Rieke, C., David, P., & Elger, C. E. (2003). Seizure prediction by nonlinear EEG analysis. *IEEE Engineering in Medicine and Biology*, 22, 57–63.
- Miletics, E., & Molnárka, G. (2004). Taylor series method with numerical derivatives for initial value problems. *J. Comp. Methods Sci. Eng.*, 4(1,2), 105–114.
- Mormann, F., Andrzejak, R. G., Egele, C. E., & Lehnertz, K. (2007). Seizure prediction: The long and winding road. *Brain*, 130, 314–333.
- Netoff, T. I., & Schiff, S. J. (2002). Decreased neuronal synchronization during experimental seizures. *Journal of Neuroscience*, 22(16), 7297–7307.
- Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain*. New York: Oxford University Press.
- Packard, N. H., Crutchfield, J. P., Farmer, J. D., & Shaw, R. S. (1980). Geometry from a time series. *Phys. Rev. Lett.*, 45, 712.
- Pijn, J. P. M., Velis, D. N., van der Heyden, M. J., DeGoede, J., van Veelen, C. W., & da Silva, F. H. L. (1997). Nonlinear dynamics of epileptic seizures on basis of intracranial EEG recordings. *Brain Topography*, 9, 249–270.
- Press, W., Flannery, B., Teukolsky, S., & Vetterling, W. (1990). *Numerical recipes in C*. New York: Cambridge University Press.
- Quyen, M. L. V., Martinerie, J., Adam, C., & Varela, F. J. (1997). Unstable periodic orbits in human epileptic activity. *Physical Review E*, 56, 3401–3411.
- Sauer, T., Yorke, J. A., & Casdagli, M. (1991). Embedology. *Journal of Statistical Physics*, 65, 579.
- Schiff, S. J., Colella, D., Jacyna, G. M., Hughes, E., Creekmore, J. W., Marshall, A., ... Weinstein, S. R. (2000). Brain chirps: Spectrographic signatures of epileptic seizures. *Clinical Neurophysiology*, 111, 953–958.
- Schiff, S. J., Sauer, T., Kumar, R., & Weinstein, S. L. (2005). Neuronal spatiotemporal pattern discrimination: The dynamical evolution of seizures. *NeuroImage*, 28(4), 1043–1055.
- Smith, E., Liou, J.-y., Davis, T. S., Merricks, E. M., Kellis, S. S., Weiss, S. A., ... Schevon, C. A. (2016). The ictal wavefront is the spatiotemporal source of discharges during spontaneous human seizures. *Nature Communications*, 7, 11098.
- Stacey, W., Quyen, M. L. V., Mormann, F., & Schulze-Bonhage, A. (2011). What is the present-day EEG evidence for a preictal state? *Epilepsy Research*, 97, 243–251.
- Takens, F. (1981). Detecting strange attractors in turbulence. In D. A. Rand & L.-S. Young (Eds.), *Dynamical systems and turbulence* (pp. 366–381). Berlin: Springer.

- Tenenbaum, J. B., Silva, V. d., & Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, *290*(5500), 2319–2323.
- Truccolo, W., Ahmed, O. J., Harrison, M. T., Eskandar, E. N., Cosgrove, G. R., Madsen, J. R., ... Cash, S. S. (2014). Neuronal ensemble synchrony during human focal seizures. *Journal of Neuroscience*, *34*(30), 9927–9944.
- Truccolo, W., Donoghue, J. A., Hochberg, L. R., Eskandar, E. N., Madsen, J. R., & Anderson, W. S. (2011). Single-neuron dynamics in human focal epilepsy. *Nature Neuroscience*, *14*(5), 635–641.
- Tucker, L. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika*, *31*(3), 279.
- Uhlhaas, P. J., & Singer, W. (2006). Neural synchrony in brain disorders: Relevance for cognitive dysfunctions and pathophysiology. *Neuron*, *52*(1), 155–168.
- Velazquez, J. L. P., Cortez, M. A., Snead III, O. C., & Wennberg, R. (2003). Dynamical regimes underlying epileptiform events: Role of instabilities and bifurcations in brain activity. *Physica D*, *186*, 205–220.
- Whitney, (1936). Differentiable manifolds. *Ann. Math.*, *37*, 645–680.

Received September 10, 2015; accepted June 9, 2017.