Journal of Computational Neuroscience (2019) 46:91–106
https://doi.org/10.1007/s10827-018-0700-1

Predicting state transitions in brain dynamics through spectral difference of phase-space graphs

Patrick Luckett[1] · **Elena Pavelescu**[2] · **Todd McDonald**[3] · **Lee Hively**[4] · **Juan Ochoa**[5]

Received: 8 January 2018 / Revised: 3 September 2018 / Accepted: 1 October 2018 / Published online: 12 October 2018

Abstract

Networks are naturally occurring phenomena that are studied across many disciplines. The topological features of a network can provide insight into the dynamics of a system as it evolves, and can be used to predict changes in state. The brain is a complex network whose temporal and spatial behavior can be measured using electroencephalography (EEG). This data can be reconstructed to form a family of graphs that represent the state of the brain over time, and the evolution of these graphs can be used to predict changes in brain states, such as the transition from preictal to ictal in patients with epilepsy. This research proposes objective indications of seizure onset observed from minimally invasive scalp EEG. The approach considers the brain as a complex nonlinear dynamical system whose state can be derived through time-delay embedding of the EEG data and characterized to determine change in brain dynamics related to the preictal state. This method targets phase-space graph spectra as biomarkers for seizure prediction, correlates historical degrees of change in spectra, and makes accurate prediction of seizure onset. A significant trend of normalized dissimilarity over time indicates a departure from the norm, and thus a change in state. Our methods show high sensitivity (90–100%) and specificity (90%) on 241 h of scalp EEG training data, and sensitivity and specificity of 70%–90% on test data. Moreover, the algorithm was capable of processing 12.7 min of data per second on an Intel Core i3 CPU in Matlab, showing that real-time analysis is viable.

Keywords Graph spectra · Epilepsy · Seizure prediction · Phase-space graph analysis

1 Introduction

Most real world systems are complex piecewise nonlinear systems that exhibit continuous or discrete dynamic behavior (Yang et al. 2015). These systems are composed of heterogeneous subsystems that can each have their own unique distribution, features, and noise. The data collected from these systems that describe their behavior often contains many features, and exist in a high dimensional space. The brain can be viewed as such a system. The brain is believed to be a self-organizing spatially embedded network displaying self-similarities at different spatial and temporal scales (Zappasodi et al. 2014). The brain can be divided into subsystems (lobes) and sub-subsystems (areas) that each play a unique biological role in our body and consciousness.

Epilepsy is a chronic disorder characterized by recurrent seizures, which is considered a change in brain state (preictal, ictal, interictal). Seizures may vary from a brief lapse in attention to severe convulsions. Approximately 50 million people worldwide have epilepsy, making it one of the most common neurological disorders (World Health Organization 2014). While some patients respond well to medication or surgery, treatments are ineffective for roughly 25% of patients (World Health Organization 2014). Those

Action Editor: Sridevi Sarma

✉ Patrick Luckett
luckett.patrick@wustl.edu

[1] Department of Neurology, Washington University, St. Louis, MO, USA

[2] Department of Mathematics and Statistics, University of South Alabama, Mobile, AL 36688, USA

[3] School of Computing, University of South Alabama, Mobile, AL 36688, USA

[4] Oak Ridge National Laboratory (retired), Oak Ridge, TN 37830, USA

[5] Department of Neurology, University of South Alabama, Mobile, AL 36604, USA

with epilepsy are at increased risk of developing many physical and mental conditions, including depression, memory problems, heart disease, stroke, arthritis, asthma, and cancer (Epilepsy Foundation of Michigan 2011). Prolonged seizure can evolve into status epilepticus, which can lead to injury or death.

Epileptic seizure prediction is defined as the identification of a time when seizure may soon occur without prior knowledge of the exact time when it will occur (Viglione and Walsh 1975). Methods for seizure prediction include relative spectral features, frequency domain and time frequency analysis, state similarity analysis, spike rate analysis, spatiotemporal correlation, phase synchronization, time domain, and many other methods based on machine learning, statistics, and nonlinear methods (Gadhoumi et al. 2016). Epileptic seizure prediction is considered one of the most important and challenging problems in biomedical sciences worldwide (Namazi et al. 2016).

This research proposes a method of seizure prediction based on dynamical system theory, delay embedding theorems, and analysis of phase space graph spectra. Specifically, we reconstruct the phase-space of time serial scalp EEG data by applying Takens' time delay embedding theorem (Takens 1981) to form state vectors in the phase space. The multidimensional vectors represent a specific state, and the succession of vectors in the phase-space represent the evolution of the system over time. Together, these form a graph with specific spectral features which we extract and analyze to make reliable seizure predictions (between 1 and 7 h prior to onset).

The contribution of this paper is the identification of 2 novel biomarkers, namely the adjacency and laplacian spectrum of phase-space graphs, which can accurately predict a seizure from at least 1 h before onset. Our statistical results indicate both methods to be reliable and efficient seizure prediction algorithms, with sensitivity of 90–100%, specificity of 90–95%, and an accuracy of 93–97% on training data, and a sensitivity and specificity of 70–90% on test data.

This paper is organized as follows. First, we provide the necessary background for the theoretical aspects of the paper. We then provide the methodology for the research, followed by the results of the analysis. Then we provide a discussion on the results, where we compare both methods, as well as provide insight into the nature of the results and parameters. We end the paper with the conclusion and future work.

2 Background

The following subsections describe the theoretical aspects of this research. We provide a brief description of dynamical systems and phase-space analysis, followed by epilepsy and seizure prediction. The section is concluded with the aspects of graph theory used for our analysis.

2.1 Dynamical systems and phase space

The brain is considered a complex nonlinear dynamical system (Andrzejak et al. 2001). Complex systems are composed of networks containing multiple components that interact in a nonlinear manner. The two key ideas that define a complex system are emergence and self organization. Emergence is the relationship between the properties and characteristics of the system at different scales, and self organization is when the system dynamically forms complex structures and behaviors over time (Sayama 2015). Dynamical systems evolve in time, and are often defined mathematically by a coupled set of first-order autonomous differential equations (Henry et al. 2012). Data gathered from a dynamical system is often in the form of a time series. Nonlinear time series analysis involves a variety of measures that allow the identification and extraction of different characteristics of a dynamical system (Kantz and Schreiber 2004).

Phase-space analysis is a method used for nonlinear time series. The phase space of a dynamical system is a representation of all possible states of a system, where each state corresponds to a unique point in the phase space (Henry et al. 2012). Every parameter of the system represents an axis of the multidimensional space. Phase space reconstruction involves taking a uniform sampled time series and applying a time delay embedding theorem to form state vectors in the phase space. The multidimensional vectors represent a specific state, and the succession of vectors in the phase space represent the evolution of the system over time.

Takens' time delay embedding theorem (Takens 1981) provides the conditions under which a smooth attractor can be reconstructed from a set of observations, and this reconstruction preserves the topological characteristics of the attractor in time delay embedded space (Kannathal et al. 2006). A dynamical system's state at a point in time t can be described by a point χ contained in $\mathbb{R}^m$. The system evolves in time according to a map $\chi(t_0) \rightarrow \chi(t)$. A simple example is a system that is fully described by a set of known differential equations. In our case the dynamical system is the brain. However, unlike a system fully described by a set of equations, the brain must be viewed as a black box. This is because we can not say with confidence what the current state of the brain is, nor can we define the map that describes the evolution of the system. This is due, in part, by the chaotic nature of brain dynamics (Carney et al. 2011). However, we are able to observe a measurable function $\phi(\chi)$ describing the state of the system via some property of the system where $\phi : \mathbb{R}^m \rightarrow \mathbb{R}$, and the system's evolution

is now described by the time series ϕ_n produced by $\phi(\chi(t))$. This is the EEG signal. Takens' theorem states there exists an embedding dimension δ and a time delay τ which can map ϕ to phase-space vectors $y_i = [\phi_i, \phi_{i+\tau}, ..., \phi_{i+(\delta-1)\tau}]$ which become deterministic at a finite dimension, and can reconstruct the state space and dynamics of the original system.

2.2 Seizure prediction

Epilepsy treatment is presently based on anti-seizure medication, which attempts to manage the seizures rather than the disease; these medications are ineffective for many patients (up to 30%) (Pauletti et al. 2017). Unfortunately, a treatment for modifying epilepsy itself does not exist. This inadequacy necessitates the need for identifying methods capable of predicting seizures prior to onset.

From a data science perspective, seizures can be viewed as a type of anomaly in the brain. As defined by Chandola et al. (2009), in a collective anomaly an individual instance may not be anomalous, but the collection as a whole is anomalous. Collective anomalies occur in a group over many consecutive time intervals (Zheng et al. 2015), and may suggest the underlying problem is embedded in a complex heterogeneous system. Seizures, which are temporary alterations in brain dynamics because of excessive neuronal activity and synchronization among groups of neurons (Kandel et al. 2000), can thus be viewed as a biological collective anomaly. Focal seizures originate in a localized and relatively small group of neurons within the brain, while general seizures can encompass a much larger area (Kandel et al. 2000).

It is believed that seizures are not abrupt events, but prior to onset are preceded by a buildup period, which can begin hours to days before the event (Litt et al. 2001; Badawy et al. 2009). Researchers (Osorio et al. 2009; Cook et al. 2014) have concluded that past events have an impact on both the timing and behavior of future events over periods ranging from days to months, and research (Sackellares 2008) suggests seizures are preceded by characteristic changes that are detectable minutes before onset. Even for individual neurons located outside the epileptic focus, research (Truccolo et al. 2011) has shown that changes in neuronal spiking activity minutes before seizure onset deviates from the norm. Moreover, the deviations were not consistent. Such heterogeneity in spiking rate suggest the underlying cause of seizure onset may be much more complex than uncontrolled mass excitation leading to a hyper-synchronous state.

A number of methods have been proposed for seizure prediction. Williamson et al. (2012) analyzed the principal components of the eigenspectra of space delay correlation and covariance matrices from EEG data at different scales using a support vector machine (SVM) classifier. Their method achieved an AUC of .973. Martis et al. (2013) used a combination of intrinsic time scale decomposition and decision tree classifiers to achieve 95.67% accuracy in seizure prediction with 10 fold cross validation. Zheng et al. (2014) proposed a method of seizure prediction using the phase synchronization information of EEG signals. The instantaneous phase of the intracranial EEG was identified using bivariate empirical mode decomposition and Hilbert transformation, and used to calculate the mean phase coherence of phase coupling strength between different EEG channels. An increase and decrease of phase synchronization occurred prior to seizure onset, and was used for classification. These and many other methods are listed in Table 1.

Many challenges exist for seizure prediction. First, there is the choice of using scalp versus intracranial EEG. Scalp electrodes placed on individuals record field potentials of ensembles of neurons (Kandel et al. 2000). However, scalp EEG reflect the activity of cortical neurons near the electrode (Osorio et al. 2016) and therefore cannot measure areas deep within the brain, such as the mesial temporal regions or thalamus. Scalp EEG is also heavily contaminated by distortions in the recorded signal caused by muscle movements and environmental factors (Hively et al. 2013). On the other hand, intracranial EEG are invasive and can lead to severe complications. Cook et al. (2013) performed the first long term study of ambulatory EEG monitoring using intracranial EEG. They placed two electrode arrays with 16 platinum iridium contacts in silicone carriers on the cortical surface of 15 patients with epilepsy, and within 4 months there were 11 adverse events of which two were considered serious. Another challenge is the nature of the data. EEG data has been shown to be both non-stationary and nonlinear, typically having at least one positive Lyapunov exponent, which is the hallmark of chaos (Carney et al. 2011). This makes typical data analysis techniques problematic.

Mormann et al. (2006, 2007) proposed guidelines for seizure prediction studies. Among these are statistical validation, event detection, implementation on a device suitable for ambulatory monitoring, and the algorithm should not be patient specific. While a "one size fits all" algorithm would be ideal, recent work (Freestone et al. 2017; Karoly et al. 2016) suggest biomarkers of seizures may be more patient specific than once thought.

2.3 Graph theory

The spectrum of a graph has been widely used in graph theory to identify the properties, characteristics, and structure of a graph (Wilson and Zhu 2008). A number of matrix representations of graphs can be used to

Table 1 Seizure prediction summary

Author	Recording	Method	Data set	Result
Zandi et al. (2010)	S	Zero Crossing Rate	561 h	Sensitivity: 88.34%
Williamson et al. (2012)	I	Eigenspectra/SVM	21	AUC: .973%
Acharya et al. (2012)	S/I	Nonlinear Features/Fuzzy Classifier	300	Accuracy: 99.7%
Martis et al. (2013)	S/I	Time Scale Decomposition/Decision Tree	60	Accuracy: 95.67%
Li et al. (2013)	S	Spiking Rate	21	Sensitivity: 75.8%
Cook et al. (2013)	I	Unknown	15	Accuracy: 65-100%
Wang et al. (2013)	I	K Nearest Neighbor	10	Sensitivity: 73%
Zheng et al. (2014)	I	Phase Synchronization	21	.5 FP/H
Ghaderyan et al. (2014)	I	Linear Features/K Nearest Neighbor	18	Sensitivity: 100%
Vahabi et al. (2015)	I	Time Frequency Analysis/Correlation Measures	21	Sensitivity: 100%
Bandarabadi et al. (2015)	S/I	SVM	3565 h	Sensitivity: 75.8%
Gadhoumi et al. (2015)	I	Cumulants/State Similarity Measures	17	Sensitivity: 80.5%
Namazi et al. (2016)	S	Hurst Exponent	120	Accuracy: 100%
Chu et al. (2017)	S	Fourier Coefficients	583 h	Sensitivity: 86.67%
Meghdadi et al. (2017)	S	Nonlinear Smoothness Analysis	1	POC
Xiao et al. (2017)	I	Naive Bayes, LDA	10	Accuracy: 82%
Kiral-Kornek et al. (2017)	I	Deep Learning	10	Sensitivity: 69%
Ibrahim et al. (2017)	S	Shannon Entropy, KNN	10 (570h)	Accuracy: 76%
Yoo (2017)	I	Power Spectral Densities, SVM	10	Accuracy: 80%
Cho et al. (2017)	S	Phase Locking Value, SVM	65	Accuracy: 83%

S = scalp recording, I = intracranial recording, POC = proof of concept, LDA = linear discriminant analysis, SVM = support vector machine, KNN = K nearest neighbor

extract spectral information, including the adjacency and the Laplacian matrices. The eigenvalues of an $n \times n$ matrix A are the roots of its characteristic polynomial,

$$p(\lambda) = \det(A - \lambda I_n). \tag{1}$$

Counting multiplicities, an $n \times n$ matrix A has n eigenvalues. The spectrum of a graph is the set of eigenvalues of the associated $n \times n$ matrix A:

$$S_A = (\lambda_1, \lambda_2, \ldots, \lambda_n), \text{ with } \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n. \tag{2}$$

The choice of matrix has a significant impact on the suitability of spectrum for a given task. In this research we evaluate the adjacency and laplacian spectrum.

Consider a graph G without multiple edges whose set of vertices is $\{v_1, v_2, \ldots, v_n\}$. The adjacency matrix $A = [A_{ij}]$ is the $n \times n$ matrix given by $A_{ij} = 1$, if there is an edge between v_i and v_j and $A_{ij} = 0$, otherwise. Our analysis assumes a loop free graph, thus $A_{ii} = 0$, for $1 \leq i \leq n$. The adjacency matrix A is symmetric with real entries, and thus its eigenvalues are all real (Fraleigh et al. 1995). Moreover, $Tr(A)$, the sum of all diagonal entries, is zero.

The spectrum of the adjacency matrix has a unique feature. For the $n \times n$ adjacency matrix A, we write the coefficient of λ^{n-1} of the characteristic polynomial of A,

$$p(\lambda) = \det(A - \lambda I_n) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n), \tag{3}$$

in two ways

$$(-1)^{n-1}Tr(A) = (-1)^{n+1}(\lambda_1 + \lambda_2 + \ldots + \lambda_n), \tag{4}$$

to obtain

$$Tr(A) = \lambda_1 + \lambda_2 + \ldots + \lambda_n = 0. \tag{5}$$

This implies that, with the exception of graphs which are completely disconnected, the adjacency spectrum contains both positive and negative eigenvalues. For example, the complete graph on n vertices K_n has adjacency spectrum $S_A = (n-1, -1, \ldots, -1)$, where -1 has multiplicity $n-1$. And the complete bipartite graph $K_{n,m}$ has adjacency spectrum $S_A = (\sqrt{mn}, 0, \ldots, 0, -\sqrt{mn})$, where 0 has multiplicity $m + n - 2$. The reason for highlighting this feature is described in the methodology. The range of eigenvalues is governed by the following inequalities:

1. The largest eigenvalue of G, λ_1, satisfies the inequalities

$$d_{ave} \leq \lambda_1 \leq d_{max} \tag{6}$$

where d_{ave} and d_{max} represent the average degree and the maximal degree of a vertex of G, respectively (Bollobás 2013; Brouwer and Haemers 2012).
2. By the Perron-Frobenius theorem we also know that all other eigenvalues λ of G satisfy $|\lambda| < \lambda_1$ (Gantmacher 1960).

The Laplacian matrix of a graph G is the matrix $L = D - A$, where A is the adjacency matrix of G and D is the diagonal matrix whose $(i, i)-$entry equals the degree of the vertex v_i. The Laplacian matrix is symmetric with real entries and therefore has all real eigenvalues. Since the entries in each row of L add to zero, $\det(L) = 0$ and therefore zero is an eigenvalue of L. Moreover, the Laplacian matrix is positive-semidefinite, thus all its eigenvalues are non-negative.

3 Methodology

Our analysis uses time serial EEG data in the 10–20 system from two scalp electrodes (F8 and FP2) to form a single channel of data in the bipolar montage, F8-FP2. The choice of F8-FP2 is 2 fold:

1. In previous work (Hively et al. 2005) we analyzed many combinations of EEG channels in the 10–20 system and the combination F8-FP2 consistently outperformed other combinations in event forewarning;
2. Based on our previous work (Ashbee et al. 2014; Hively et al. 2013), we believe these electrodes, located on the right-frontal region, act as a filter for pre-ictal condition change.

The data were acquired in epilepsy monitoring units (clinical setting) under standard clinical protocols from 41 temporal lobe epilepsy patients. The archival data were provided only after the fact under appropriate Institutional Review Board (IRB) guidelines, namely all personal identifiers were removed. Ages ranged from 4 to 57 years old. Of all data sets, 36 were from females and 24 were from males. The data sets range in length from 1.4 to 8.2 h with an average of 4.4 h. Forty data sets had seizures, and twenty did not. (Non)event determination in the datasets was made by neurologists who were board certified in epilepsy. No filter was applied to the data when collected. Data were evaluated for quality following methods proposed in (Hively and Ng 1998; Hively 2009). This method verifies several features in the data, including evaluation of proper number of data points, adequate sampling rate, and lack of excessive periodic content. Data sets that failed one or more of these tests were rejected and not considered in our analysis. However, all data sets passed all of the quality tests. Further details on this data can be found in Table 2 (in no particular order).

The data are uniformly sampled in time, t_i, at 250 Hz. The data are divided into equal size cutsets, resulting in N time-serial points for each cutset, $e_i = e(t_i)$. The cutsets are used to characterize the behavior of the system over a particular time segment. Possible values for N range from 10,000 to 100,000.

A zero-phase, quadratic filter removes electrical artifacts which would otherwise obscure the event. The filter retains the nonlinear amplitude and phase information of the original e_i data (Hively et al. 1995). The filter uses a moving window of $2w + 1$ points of $e(t_i)$ data, which are fitted to a parabola in a least-squares sense, yielding $N - 2w$ points of artifact data, f_i. Here, w is half the width of the sliding window, and is a trainable parameter. The residual signal has essentially no low-frequency artifacts, $g_i = e_i - f_i$.

After filtering, every data point (g_i) is converted to a symbolized form (s_i). Symbolization is a binning strategy used to balance the need for precise data, and the need to exclude as much signal noise as possible. The number of symbols is predetermined such that all data points in the filtered cutset becomes one of S different integers $0, 1, ..., S - 1$ based on the range of the data using the following formula:

$$0 \leq s_i = INT\left[S\frac{g_i - g_{min}}{g_{max} - g_{min}}\right] \leq S - 1, \tag{7}$$

where INT rounds the calculated value down to the next lowest integer, and g_{max} and g_{min} represent the largest and smallest values of g_i respectively. For each data set being analyzed, g_{max} and g_{min} correspond to the minimum and maximum value that occur in the first cutset of data for that data set. In general, this does not correspond to the minimum and maximum that occur in the data set as a whole, but still allows a uniform scaling of all data points.

Taken's Theorem (1981) guarantees that, with a time-delay-embedding of a sufficiently high dimension, the features of topology in a dynamical system can be reconstructed from a finite data set. A time-delay vector converts the symbolized data into a series of unique, delay-embedded phase-space vectors y_i:

$$y_i = [s_i, s_{i+L}, ..., s_{i+(D-1)L}] \tag{8}$$

Here, L represents the amount of time delay (Lag), which must not be too small (making s_i and s_{i+L} indistinguishable) or too large (making s_i and s_{i+L} independent by long-time unpredictability). The embedding dimension of the phase-space vectors is D, which must be sufficiently

Table 2 Details of training data

Patient	Data set	Sex	Age	Seizure	Seizure type	Activity prior to seizure
3	12	F	15	Y	partial	–
4	13	F	44	Y	partial	reading
5	16	M	7	Y	partial	TV, eating
99	17	F	17	N	–	–
6	18	M	43	Y	generalized	asleep
6	19	M	43	Y	generalized	asleep
100	22	M	21	N	–	–
8	24	M	30	Y	partial/generalized	lying in bed, talking
8	26	M	30	N	–	–
11	37	F	42	N	–	–
11	39	F	42	Y	partial	lying in bed, reading
11	125	F	42	Y	partial/generalized	Unknown
11	131	F	42	Y	partial	TV, lots of Movement
12	42	M	15	Y	partial	eating, talking
12	46	M	15	Y	partial	talking, sitting in bed
19	127	M	4	N	–	–
19	129	M	4	Y	partial/generalized	awake
22	149	F	43	N	–	–
22	150	F	43	N	–	–
24	157	F	41	N	–	–
24	158	F	41	N	–	–
24	163	F	41	N	–	–
24	165	F	41	N	–	–
27	170	M	37	Y	partial	–
75	193	M	41	Y	partial	–
81	199	M	41	Y	partial	unknown
82	200	M	23	Y	generalized	sitting in bed, playing
84	203	F	33	Y	partial/generalized	asleep
83	207	F	33	Y	partial	asleep
86	211	M	41	Y	partial/generalized	asleep
74	214	F	32	Y	partial	awake, TV
74	216	F	32	Y	partial	awake, lying in bed
89	222	F	52	Y	partial	sitting in bed, talking
90	221	M	43	Y	partial	sitting in bed, eating
40	235	F	51	Y	partial/generalized	sitting in bed, talking
46	255	F	16	Y	partial	–
47	259	F	43	Y	partial	lying in bed
48	261	F	14	N	–	–
61	264	M	33	Y	partial	–
61	265	M	33	Y	partial	asleep
61	266	M	33	Y	partial	sitting in bed
61	267	M	33	Y	partial	asleep
62	270	F	27	N	–	–
61	271	M	33	Y	partial	talking, sitting in bed
63	273	F	44	N	–	–
64	274	F	56	N	–	–
64	275	F	56	N	–	–
64	276	F	56	N	–	–

Table 2 (continued)

Patient	Data set	Sex	Age	Seizure	Seizure type	Activity prior to seizure
67	283	F	20	N	–	–
67	285	F	20	N	–	–
65	284	M	58	N	–	–
68	286	M	56	Y	partial/generalized	asleep
66	287	M	19	Y	partial	talking
69	289	F	57	Y	partial	asleep
69	299	F	15	Y	partial	asleep
70	293	F	44	Y	partial	lying in bed, writing
72	300	F	44	Y	partial	asleep
94	308	F	15	Y	partial	–
30	386	F	15	Y	partial	–
54	403	M	12	Y	partial	–

large to capture the dynamics, but not too large to avoid over-fitting. The phase-space vectors, y_i, form phase-space graph nodes in the phase space which is a subset of the $D-$dimensional Euclidean space.

The variable M defines a second set of time-delay states which represent the process flow $y_i \longleftrightarrow y_{i+M}$. This process flow defines local state transitions within a cutset, and represents the undirected edges connecting phase-space graph nodes. The nodes (y_i) and links ($y_i \longleftrightarrow y_{i+M}$) form a graph G:

$$G = (Y, B) \tag{9}$$

where vertices $Y = \{y_1, y_2, ..., y_n\}$ represents the set of phase-space nodes, and $B = \{b_1, b_2, ..., b_m\}$ is the set of links (edges) whose elements are two element subsets of Y with topologically-invariant measures that are independent of any unique labeling. These graphs are undirected, unweighted, and simple. A list of all the previously mentioned parameters, their range, and the effect of choosing certain values are listed in Table 3. The algorithm uses dissimilarity measures between graphs generated in different cutsets to identify state transitions. The dissimilarity measure used in this research is the spectral distance of the phase-space graph adjacency and laplacian matrix.

The spectral distance (SD) of two graphs with n vertices each is defined as the Euclidean distance between the two graph spectra (Wilson and Zhu 2008):

$$SD(G_1, G_2) = \sqrt{\sum_{i=1}^{n} (\lambda_{(1)i} - \lambda_{(2)i})^2} \tag{10}$$

Based on the parameters used to construct the phase space graph, we know before hand the maximum number of possible nodes that could be created:

$$\text{Number of Nodes} = S^D, \tag{11}$$

where S is the number of symbols and D the number of dimensions. However, we do not know the actual number of nodes that will be created. In fact, we have observed in practice that it is common to have significantly less nodes created compared to the number of nodes possible. This attribute brings to light the important fact that the size of graphs produced in each cutset varies, even for the same parameter set. This is important to consider because the number of spectral values is equal to the number of nodes in the graph. As such, the spectral distance of different size graphs becomes less intuitive. As previously mentioned the sum of the eigenvalues of the adjacency matrix is equal to 0, and therefore the spectrum contains both positive and negative values around the zero average. Intuitively, we would want the positive values of one graph to be compared to the positive values of another, and the same for negative values.

The common approach to overcome the issue of graphs of different sizes is to pad the smaller graph with extra vertices to make the adjacency matrices, and thus the spectra, the same size. As these extra vertices are disconnected from the original components of the graph, the effect on the smaller spectrum is the addition of zeros. This method is employed in our analysis when comparing the adjacency and laplacian spectra of G_1 and G_2.

The eigenvalues of the given matrix are calculated for each cutset of data. Initially, these values are calculated for the first B (number of base cases) cutsets, which are known not to contain event data. Then, each base case is compared to every other base case, which yields spectral distances $SD(i)$, for $1 \leq i \leq \binom{B}{2}$. We then calculate the mean spectral distance (MSD):

$$MSD = \frac{\sum_{i=1}^{\binom{B}{2}} SD(i)}{\binom{B}{2}}, \tag{12}$$

Table 3 Summary of trainable parameters

Parameter	Description	Low Value	High Value	Range
BC	number of base cases	short baseline	long baseline	5–15
D	dimension	under fitting	over fitting	2–9
K	successive occurrences	short prediction	long prediction	1–50
L	time delay (lag)	small unfolding	excessive unfolding	10–80
M	state transitions	short correlation	long correlation	10–80
N	points per cutset	scarce statistics	blurred change	10,000–100,000
S	number of symbols	noise rejection	too precise	2–9
U	threshold	small change	large change	(−5)–5
w	filter half width	fast artifact	slow artifact	10–100

followed by the spectral standard deviation (SSD):

$$SSD = \sqrt{\frac{\sum_{i=1}^{\binom{B}{2}} (SD(i) - MSD)^2}{\binom{B}{2} - 1}}. \tag{13}$$

Following these calculations, every test case spectral distance (TCD) is calculated by averaging the difference of the test case with every base case, and then normalizing that value with the MSD and SSD to produce a single normalized dissimilarity value (spectral distance) for the given test case (TC).

$$TCD = \frac{\left(\frac{\sum_{i=1}^{B} SD(TC, BC_{(i)})}{B}\right) - MSD)}{SSD}. \tag{14}$$

The dissimilarity measure value defines, in relation to a predefined threshold, the degree to which a given test cutset varies from the baseline. Exceeding the threshold dictates a tabulation response from the algorithm; if K successive threshold crossings are recorded, an event is indicated by the algorithm. Both the size of the threshold and the number of successive occurrences above the threshold indicate an anomaly. Both are trainable parameters which are optimized for any given parameter set. The process of monitoring a patient by acquiring a single cutset of two channels of raw EEG data, filtering and combining the 2 channels, constructing the phase space graphs, calculating and comparing the spectrum, and taking the appropriate action is summarized in Fig. 1.

Many measures are available for statistical evaluation. One measure is the number of true positives (TP) for known event data sets (Ev), to yield the true positive rate (sensitivity) of TP/Ev. A second measure is the number of true negatives (TN) for known non-event data sets (NEv). The true negative rate is TN/NEv (specificity). The goal is a sensitivity and specificity of unity. We seek to minimize the prediction distance:

$$PD = \left\{ \left[1 - \frac{TP}{Ev}\right]^2 + \left[1 - \frac{TN}{NEv}\right]^2 \right\}^{1/2}, \tag{15}$$

which is an appropriate objective function for any event.

A computationally-intensive training period can select the best phase-space parameters through repeated Monte Carlo runs. Decreased prediction distance value indicates greater efficacy for a set of parameter values. Random selection of phase-space parameters is efficient because the resulting prediction distance exhibits markedly fractal behavior, and thus cannot reliably be chosen with a genetic or statistical algorithm.

A requirement of any statistical test of significance is an optimal data set. A small data set produces an imprecise estimate of accuracy with a wide, uninformative confidence interval, and may not capture realistic fluctuations in data. As a rule, data sets should represent the population and be sufficiently diverse. Our training data set consists of 60 characterized files from different genders, races, and age groups, which satisfy the representation and diversity requirements in the context of seizure prediction.

4 Adjacency spectrum results

To identify the optimal parameter set for eigenvalues we performed a Monte Carlo search of parameters. The best results achieved corresponded to a prediction distance of .1, an accuracy of 97% (58/60), a sensitivity of 100% (40/40), a specificity of 90% (18/20), and an AUC of .97 (Fig. 3). The mean prediction time was 3 h prior to onset with a standard deviation of 1.2 h. The median

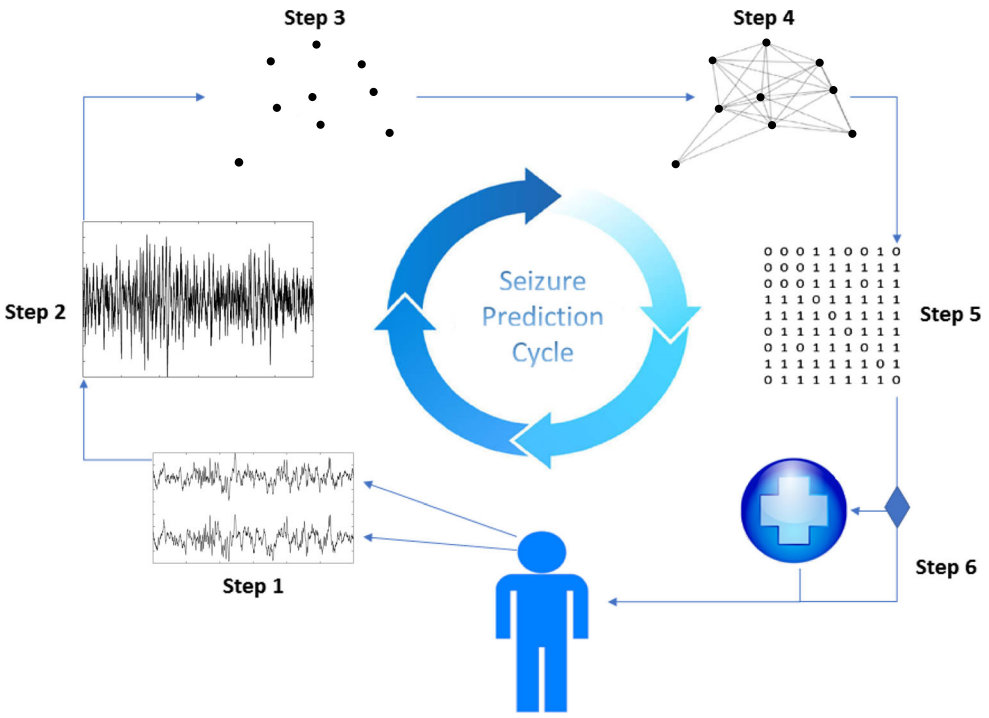

Fig. 1 Seizure prediction cycle: Each cycle represents acquiring and processing a single cutset of data. Step 1: Acquire two channels of scalp EEG from patient (F8, FP2). Step 2: Subtract the value of channel 1 (F8) from channel 2 (FP2), F8-FP2, and filter the resultant bipolar channel. Step 3: Construct D-dimensional phase space vectors via time delay embedding of symbolized data. Step 4: Construct links between state vectors to simulate process flow. Step 5: Construct adjacency or laplacian matrix from the graph produced (Step 4) and calculate eigenvalues to compare with preictal baseline. Step 6: If a significant difference is observed over a given number of cycles (successive occurrences of normalized spectral distance above a predefined threshold), notify the patient, patient caregiver, or first responders. If no significant difference exists continue process

and mode prediction time were both 3 h. The optimal number of successive occurrences of cutsets of data whose spectral difference was above a given threshold was five, corresponding to roughly 6.7 concurrent minutes. These results correspond to zero false negatives and two false positives. The parameter values for these results are: $D = 2, S = 3, L = 50, M = 46, N = 20676, BC = 9$. Figure 2 is an example of dissimilarity measures of the adjacency spectrum spectral difference in seizure and non seizure readings with the threshold marked, and Table 4 depicts the output of the algorithm on several seizure and nonseizure files.

Twenty additional data sets were reserved (not trained on) for out of sample testing. These data sets were independent of the 60 training data sets. Ten of the data sets contained seizures and ten did not. The data were acquired in epilepsy monitoring units (clinical setting) under standard clinical protocols with Institutional Review Board (IRB) approval and all personal identifiers removed. (Non)event determination in the datasets was made by neurologists who were board certified in epilepsy. No filter was applied to the data when collected, and it was sampled at 250 Hz. The data sets ranged from 1.5 to 4 h. No patient demographics were provided. Data sets containing seizures were verified to have seizure onset at least 1 h from the beginning of the reading. The previously identified parameter sets were used for testing. The test using adjacency eigenvalues achieved an accuracy of 75%, specificity of 70% (7/10), sensitivity of 80% (8/10), a false positive rate of 30% (3/10), and a false negative rate of 20% (2/10) on the test data. We further evaluated the mislabeled data sets and found that a small adjustment to the threshold (from .366 to .345) raised the sensitivity to 90% (9/10) and overall accuracy to 80% (16/20). The specificity remained at 70% (7/10), the false positive rate remained at 30% (3/10), and the false negative rate was 10% (1/10). This was achieved by leaving every parameter in the phase-space reconstruction the same, and only adjusting the threshold. The mislabeled files that were correctly relabeled after adjusting the threshold were from different patients, and these patients had only one file in the test set. This shows that the algorithm generalizes well for a given parameter set (phase-space reconstruction), but may require some fine tuning to the threshold for some patients, which is on par with current research suggesting seizure prediction may be somewhat patient specific (Freestone et al. 2017; Karoly et al. 2016).

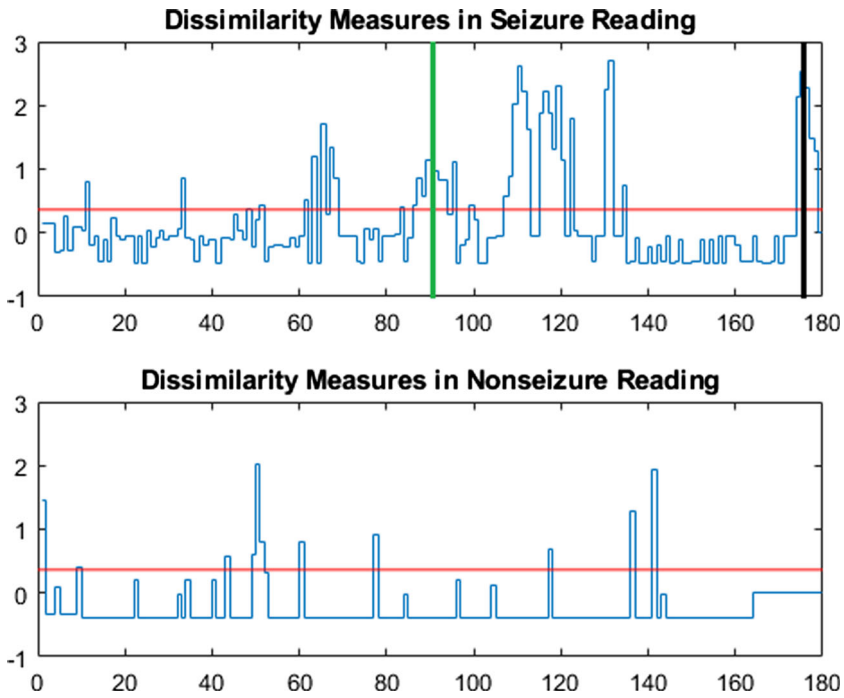

Fig. 2 Dissimilarity of normalized adjacency spectral distance in a reading containing a seizure (top), and a reading with no seizure (bottom). The red line indicates the threshold for measuring successive occurrences, the black line indicates seizure onset, and the green line indicates the first notification of a seizure by the algorithm, which was roughly 2 h prior to onset. Reading duration roughly 4 h

5 Laplacian spectrum results

After performing the Monte Carlo search of parameters using the laplacian spectrum, we achieved a prediction distance of .09, an accuracy of 93% (56/60), a sensitivity of 93% (37/40), a specificity of 95% (19/20), and AUC .94 (Fig. 3). The mean prediction time wwas 3.3 h prior to onset with a standard deviation of 1.1 h. The optimal number of successive occurrences was five. The parameter values for these results are: $D = 2$, $S = 3$, $L = 24$, $M = 65$, $N = 23057$, $BC = 10$. The same out of sample testing process was followed, and resulted in a specificity of 60% (6/10), sensitivity of 70% (7/10), a false positive rate of 40% (4/10), and a false negative rate of 30% (3/10) before threshold tuning, and a specificity and sensitivity of 80% (8/10) and false positive and false negative rate of 20% (2/10) after threshold tuning. The overall accuracy of the model was 80% (16/20). Again, the mislabeled files that were correctly relabeled were from different patients, and these patients had only one file in the test set. The MATLAB implementation of both methods can be found online (https://github.com/luckettp/Predicting-State-Transitions-in

Table 4 Tracking values for several seizure and non seizure files

File	Seizure	Tracking
1	N	10000000000000000000000000000000000000000000000000000000000000000000000000000100000000
2	N	10000000000000000000000000000000000000000000000000000000000000000000000000000100000000000000000000
3	N	10000000000000010000100000000000000000110000000000010000000101100010001000
4	N	00000100000000011100110101101100100111101101110001000111100000000100111010111
5	N	00000000000000000000000000000000000000000000000000000000000000000000000000000000000
6	N	01000000000000010000000000000000000000000000000000000000011010110110001000
7	N	11000000000000000000000000100000000001000000000000000
8	Y	000101000000011111F001101011010001010010100001111100001111111111111111111111111111111111111111111111111S—
9	Y	000000000101100000000010000011010010000010000111 1F1010111111111111111100000100000101000000111101111 1S—-
10	Y	000000001000000110000001010000000000000010000000000000000000000000010010000000000000000010111000000100S–
11	Y	10000000000001011010101111F1101101000100100001001100010000001000000000100000101101111000101111101111S—-

Column 1: File index (does not correspond to Table 1). Column 2: Y= seizure occurs in file, N = no seizure in file. Column 3: Tracking values, 0 indicates cutset is not above threshold, 1 indicates cutset dissimilarity is above threshold, F indicates seizure will occur, S indicates seizure onset

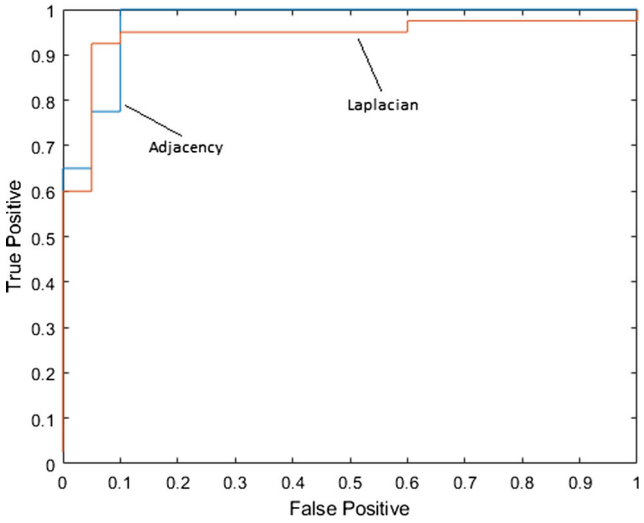

Fig. 3 ROC curves for adjacency and laplacian algorithms on training data

-Brain-Dynamics-Through-Spectral-Difference-of-Phase-Space-Graphs).

6 Discussion

The results of both analyses achieved promising results on both training and testing data. We believe this is because the spectrum of a graph is an optimal choice for phase-space graphs derived using Takens' theorem. Takens' theorem provides the topology of the underlying dynamics over a given amount of time. This topology is independent of any unique labeling and directivity. The spectrum (adjacency and Laplacian) of an undirected graph represent the structural topology of the graph, which is also independent of unique labeling. Therefore, the spectral distance between two graphs represents the difference in the dynamics as time progresses. Because a change in EEG dynamics occurs prior to seizure onset, a large spectral difference will be observed, and indicate seizure onset.

Table 1 summarizes other current research in seizure prediction. These analysis methods include traditional nonlinear measures, such as the Hurst Exponent and Fourier Coefficients, neuroscience based methods such as spike rates, and a variety of machine learning techniques. Based on the reported results, our method appears to be on par with the state of the art in seizure prediction. Moreover, our method is novel/preferable in the following ways:

1. Scalp EEG Data: Most current methods analyze intracranial EEG data. While intracranial EEG data has far less noise and is closer to the seizure source, the possibility of infections and other complications make it problematic for long term use. Other methods may use scalp EEG data, but explicitly state they do not include excessively noisy data sets in their analysis, which is not a realistic practice for clinical or ambulatory care. Our zero phase quadratic filter removes artifacts and noise while retaining the amplitude and phase of the data, which allows us to accurately evaluate scalp data.
2. Two EEG Channels: Our methods yield high accuracy while only requiring two channels of scalp EEG data. This is essential for ambulatory care, as methods that require many channels would quickly drain the power source of the device because of the demand of gathering, storing, and processing the large amount of data.
3. Method: To our knowledge, our method is the only method to use Takens' phase-space embedding theorem to reconstruct the nonlinear dynamics of the EEG data, and perform spectral graph analysis to predict state transitions.
4. Data: Our training analysis contained 60 data sets from 41 different patients, which is significantly more than what other methods have used for evaluation. We also performed out of sample testing on 20 data sets, which yielded statistically significant results.
5. Computational Complexity: Our method is computationally feasible for clinical and ambulatory care. Calculating the eigenvalues for the spectrum is computationally intense for large matrices, ranging from $O(n^2)$ to $O(n^3)$ (Demmel et al. 2007) depending on the method and type of matrix under evaluation (here n is the number of vertices). A dimension of two with three symbols corresponds to a maximum of a 9×9 matrix (3^2), facilitating fast computations. On average, the algorithm could analyze 9.4 cutsets of data per second, where each cutset corresponded to roughly 1.35 min of EEG data measured at 250 Hz. This exceeds what would be required of real time monitoring. These times were calculated using Matlab version 9.1.0.441655 (R2016b) on a Lenovo ThinkPad with an Intel Core i3-5005u CPU at 2.00 GHz and a 64 bit operating system.

Both methods showed similar results, and both methods had very similar prediction times. However, it should be noted that both methods searched for an optimal parameter set independently using different random number generators (Java 7).

We evaluated the raw data to identify a possible reason for the occurrence of false negatives. For the false negative file, we identified several extreme spikes that occurred in the first cutset. Figure 4 depicts these spikes, which are significantly larger than any other data point, and likely the result of noise caused by an environmental factor such as static electricity. Recall that the scaling values for

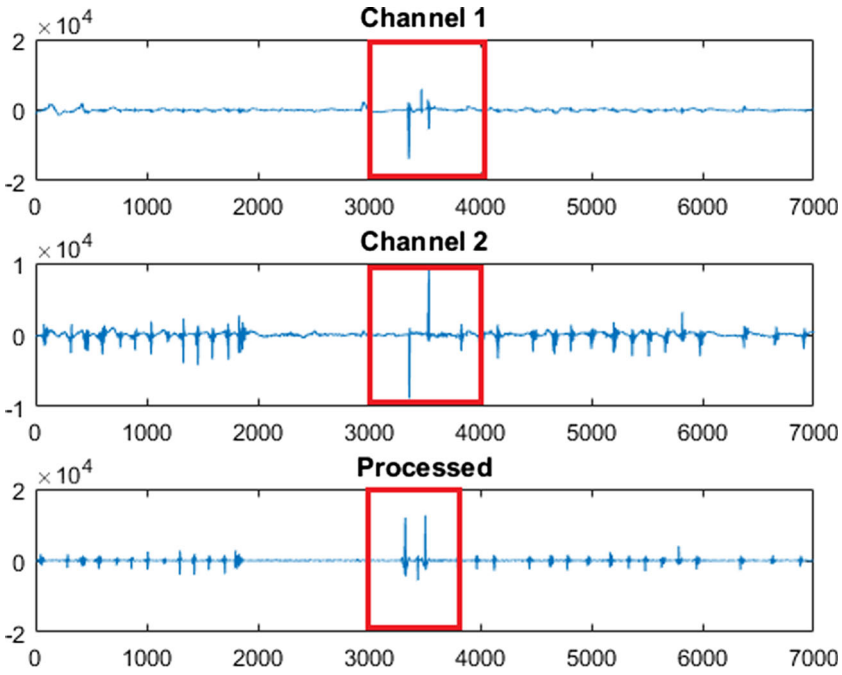

Fig. 4 Raw and filtered data from first cutset of false negative file. Because this data was at the beginning of the recording, it was treated as a base case, and compared to all test cases. Top: Channel F8 from first base case. Middle: Channel Fp2 from first base case. Bottom: Bipolar F8-Fp2 after applying our filtering strategy. Extreme outliers which were not removed by subtracting and filtering data channels could result in obscuring seizure activity, specifically if these outliers occur in base cases

symbolization are chosen from the minimum and maximum from the first cutset of data. In some cases, such outliers can have a minimizing effect on the graphs in certain cutsets. To test the hypothesis that the outliers in the first cutset may be the cause of the incorrect labeling, we evaluated the data set by scaling any value in the first cutset to the mean if it was larger than two standard deviations from the mean or smaller than negative two standard deviations from the mean. The file was retested using the same threshold and parameters, and was successfully able to predict the seizure 3 h prior to onset. However, the scaling had a deleterious effect on the results as whole. This anomaly is an indication that some fine tuning may be appropriate on a patient to patient basis.

We also considered the possibility that the age difference could cause less than optimal results. This is because children show different patterns in EEG than adults (Barriga-Paulino et al. 2011). We reanalyzed the training data by removing any patient recordings under the age of 18 and found no significant difference in the performance of the algorithm. We believe this is due to 1) Takens' theorem extracts the underlying dynamics of the data, which may be similar for adults and children, and 2) the threshold is optimized based on the set of files it is evaluating.

Another consideration is how well the algorithm generalizes to different patients, given that all patients can show slight to moderate differences in signal patterns (Huang et al. 2012). Our algorithm showed promising results on over 40 different patients, which shows the algorithm can generalize well. We believe this is because 1) our filtering strategy removes high frequency oscillations which could be patient specific, 2) symbolizing the data into a discreet set of "bins" (three in our case) further removes patterns that are unique to a patient, and 3) Takens' theorem captures the underlying dynamics of EEG data over time. Many patterns that can occur due to specific cognitive task or motor functions may not occur frequently enough to have a significant impact on the underlying dynamics. These facts are reflected by the algorithms ability to optimize a threshold with high accuracy on many different patients.

Regarding the false positives, we identified a significant number of artifacts caused by muscular activity in the cutsets where the algorithm flagged a seizure. Figure 5 is an example of these artifacts. While we do employ a filtering strategy to remove artifacts, it is reasonable to assume that if a significant amount of data in the cutset is contaminated, the results will be slightly obscured.

A number of interesting features were identified in the analysis as compared to our previous methods that used unique nodes and edges, as well as other graph characteristics and features as a measure of dissimilarity (Hively et al. 2013; Luckett et al. 2017). First, this method was less sensitive to small perturbations in parameters. We found that once a "good" parameter set was identified, we could slightly alter the link parameter M, the lag L, the number of base cases, or the size of the cutset while not affecting the results of the algorithm. The only parameters that could not be changed without altering the results were the dimension and the number of symbols. This comes as no surprise, as the number of symbols and the dimension define the number of possible states of the system. Figure 6 shows the effect of changing parameter values on dissimilarity

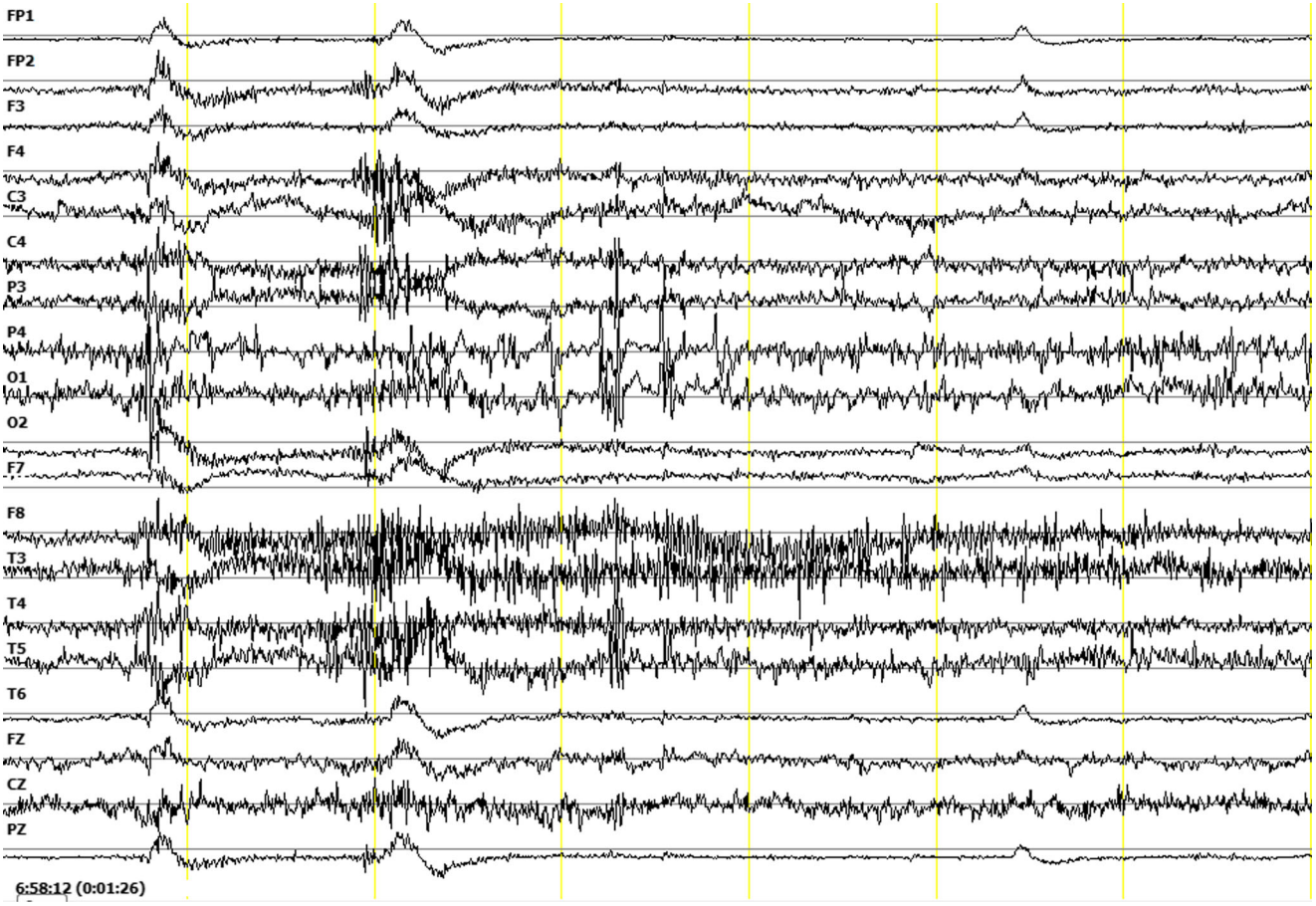

Fig. 5 Muscular artifacts: If a significant amount of muscular artifacts are observed in a window of data, the filtering strategy will have difficulty removing the artifacts because there is not a significant source of "clean" data in the window. This can lead to an indication of a seizure in a non-event file, especially if the data is in a base case

measures as compared to our baseline parameters. As an interesting aside, we analyzed the correlation between a change in phase space parameters and over 30 features of the resulting graphs, including smallest and largest eigenvalues, diameter, and algebraic connectivity and found the only parameter consistently correlated with a change in the structure and spectrum of the graph was the dimension.

The second interesting feature is that the algorithm performed well when using a smaller set of parameters, notably the dimension and number of symbols. We analyzed over 250,000 parameter sets with randomly generated parameters. Both dimension and symbols ranged between two and five. The dimension and symbols identified on this data set that performed best were a dimension of two and three symbols. Some may note the embedding dimension of two is very small. Takens (1981) found the reconstruction of the system dynamics requires the embedding dimension to be at least one plus twice the number of degrees of freedom, which can be measured as the correlation dimension, yielding an embedding dimension of twice the correlation dimension plus one. However, in practice the dimension is a trainable parameter to achieve the best prediction distance. This small value of D may arise because we are analyzing 60 data sets from 41 different subjects. It is easy enough to calculate the correlation dimension for a single data set from a single system and find a minimum embedding dimension, but it becomes less intuitive for multiple systems over very long scales. This variability in D is especially true for EEG, which has been shown to exhibit significant variability in correlation dimension based on a number of factors, ranging from eyes being open or shut, to cerebral hemorrhage, thrombosis, or infarctions (Zhu et al. 2008; Nan and Jinghua 1988).

The use of spectral features as a method for classification has been brought into question due to the possibility of cospectral graphs. It has been shown that the fraction of graphs with a cospectral mate is nondecreasing for small n, with a local maximum at $n = 10$ of 21.3% for the adjacency matrix (Haemers and Spence 2004). It is also known that the ratio of cospectral graphs to total number of possible graphs approaches 0 as $n \to \infty$. This ratio is 18.6% for graphs with nine vertices, which is the maximum number

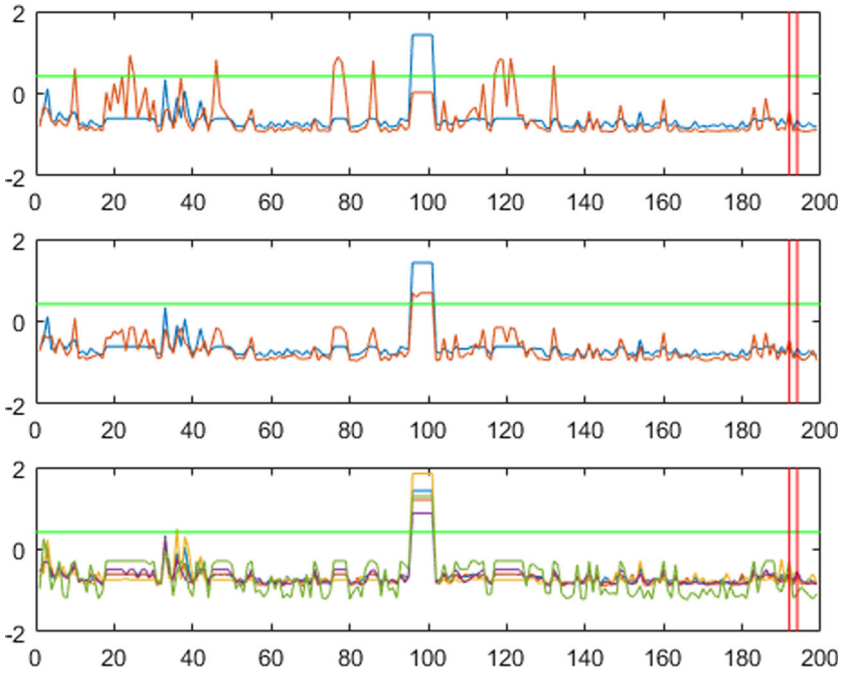

Fig. 6 Impact of changing phase space parameters on EEG data. Horizontal green line represents threshold and vertical red lines represents seizure beginning and end. Top: Blue line indicates our reported parameters and red line indicates changing dimension parameter from 2 to 3. Middle: Blue line indicates our reported parameters and red line indicates changing symbol parameter from 3 to 4. Bottom: Blue line indicates our reported parameters and other lines indicates changing lag, mu, and number of base cases. Prediction occurs at index 100, as seen in blue line (our parameters). Changing dimension and symbols shows a significant change in patterns in the data (top and middle). Changing other parameters does not introduce significant changes in data (bottom)

of vertices a graph with our parameters could have. In our analysis, the mean percent of cospectral graphs across all data sets was 53% with a standard deviation of 24%. The mean percent of cospectral graphs for nonseizure data sets was 51% and for seizure files 54%. While this is a relatively high percent, we postulate that our use of numerous base cases offset the effect of cospectral graphs in the data set. It is also reasonable to assume that in a nonseizure file with little or no epileptiform activity, the presence of cospectral graphs would indicate similarity throughout the data set, which is desirable.

7 Future work

Our future work for this research is five-fold. First, further testing of different parameter sets will be conducted to attempt to improve the results identified here. Second, we are currently in the process of annotating significantly more EEG files. Once annotated, the algorithms with the identified parameters will be tested on the new data to provide further statistical validation. Third, we intend to evaluate the performance of other distance functions, such as the Minkowski distance with $P \geq 1$, $P \neq 2$:

$$d(X, Y) = \sqrt[P]{\sum_{i=1}^{n}(x_i - y_i)^P}. \tag{16}$$

It is unknown if there is a correlation between the dimension of the phase space graph and that of the distance function (i.e. our best results were in 2 dimensions and Euclidean distance is the Minkowski distance in 2 dimensions), but it is known that on the Euclidean space $\mathbb{R}^n$ all norms are equivalent, meaning if two vectors are close in one norm, they are also close in others (up to a multiplicative constant). However, in practice we have observed changing from Euclidean distance ($P = 2$) to Minkowski distance with $P = 3$ yields different results and thresholds given the same parameters. Another issue concerns prediction horizons. This research did not consider prediction horizons when evaluating the model. Given our parameters (window size and number of successive occurrences), a prediction would have to occur at least 7 min prior to seizure onset, and as reported, the majority of predictions occur within 2–4 h of the seizure. Future work should try to identify optimal parameter tuning to achieve consistent prediction times. This work also supports our belief that the bipolar channel F8-FP2 act as a filter for pre-ictal condition change. Future work should seek to further validate this hypothesis, as well as investigate the mechanistic actions producing the results. If a specific area of the brain does act as a filter for pre-ictal change, it would have significant implications for all types of epilepsy monitoring. Lastly, once the optimal parameters and distance function have been evaluated, the algorithm will be implemented on an ambulatory device for real world testing.

8 Conclusion

The human brain is a complex networked nonlinear system that exhibits unique temporal and spatial behavior. The

brain is composed of heterogeneous subsystems that, both individually and as a group, play a pivotal role in our lives. Epilepsy is a chronic disorder characterized by excessive electrical activity. Transitions from normal to ictal states exhibit changes in brain dynamics prior to the event which can be used to predict a seizure before onset.

In this paper, we have proposed a method of seizure prediction based on spectral graph theory and phase space analysis that satisfy the goals for effective predictive biomarkers of epileptic seizures. The algorithm utilizes scalp EEG that is minimally invasive and has shown promise for high sensitivity and specificity. Reconstructing the phase-space of scalp EEG data via time delay embedding forms state vectors in the phase-space, and the succession of vectors in the phase-space represent the evolution of the system over time. The resulting graphs exhibit spectral features which we extract and analyze to make reliable seizure predictions. We performed our analysis on two spectral features. The first analysis compared the adjacency spectrum, and achieved a sensitivity of 100%, a specificity of 90%, and an accuracy of 97%, and the second analysis compared the laplacian spectrum of the graphs, and achieved an accuracy of 93%, a sensitivity of 93%, and a specificity of 95%. These methods were evaluated on 241 h of scalp EEG data from 41 temporal lobe epilepsy patients with a total of 40 data sets with seizures and 20 data sets without seizures. Further, out of sample testing was performed on both methods using twenty additional data sets. Both methods achieved statistically significant results on out of sample test, with a sensitivity and specificity of 80%–90%. The results suggest these methods are capable of accurate, reliable, and efficient predictions of state transitions in brain dynamics.

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no conflict of interest.

References

Acharya, U.R., Sree, S.V., Ang, P.C.A., Yanti, R., Suri, J.S. (2012). Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals. *International Journal of Neural Systems*, *22*(2), 1250,002.

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*(6), 061,907.

Ashbee, W.S., Hively, L., McDonald, J. (2014). Nonlinear epilepsy forewarning by support vector machines. In *Epilepsy topics*: InTech.

Badawy, R., Macdonell, R., Jackson, G., Berkovic, S. (2009). The peri-ictal state: cortical excitability changes within 24 h of a seizure. *Brain*, *132*(4), 1013–1021.

Bandarabadi, M., Teixeira, C.A., Rasekhi, J., Dourado, A. (2015). Epileptic seizure prediction using relative spectral power features. *Clinical Neurophysiology*, *126*(2), 237–248.

Barriga-Paulino, C.I., Flores, A.B., Gómez, C.M. (2011). Developmental changes in the eeg rhythms of children and young adults. *Journal of Psychophysiology*, *25*(3), 143–158.

Bollobás, B. (2013). Modern graph theory, vol. 184. Springer Science & Business Media.

Brouwer, A.E., & Haemers, W.H. (2012). *Distance-regular graphs*. Berlin: Springer.

Carney, P.R., Myers, S., Geyer, J.D. (2011). Seizure prediction: methods. *Epilepsy & Behavior*, *22*, S94–S101.

Chandola, V., Banerjee, A., Kumar, V. (2009). Anomaly detection: a survey. *ACM Computing Surveys (CSUR)*, *41*(3), 15.

Cho, D., Min, B., Kim, J., Lee, B. (2017). Eeg-based prediction of epileptic seizures using phase synchronization elicited from noise-assisted multivariate empirical mode decomposition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *25*(8), 1309–1318.

Chu, H., Chung, C.K., Jeong, W., Cho, K.H. (2017). Predicting epileptic seizures from scalp EEG based on attractor state analysis. *Computer Methods and Programs in Biomedicine*, *143*, 75–87.

Cook, M.J., O'Brien, T.J., Berkovic, S.F., Murphy, M., Morokoff, A., Fabinyi, G., D'Souza, W., Yerra, R., Archer, J., Litewka, L., et al. (2013). Prediction of seizure likelihood with a long-term, implanted seizure advisory system in patients with drug-resistant epilepsy: a first-in-man study. *The Lancet Neurology*, *12*(6), 563–571.

Cook, M.J., Varsavsky, A., Himes, D., Leyde, K., Berkovic, S.F., O'Brien, T., Mareels, I. (2014). The dynamics of the epileptic brain reveal long-memory processes. *Frontiers in Neurology*, *5*, 217.

Demmel, J., Dumitriu, I., Holtz, O. (2007). Fast linear algebra is stable. *Numerische Mathematik*, *108*(1), 59–91.

Epilepsy Foundation of Michigan (2011). http://www.epilepsymichigan.org/page.php?id=358. Website.

Fraleigh, J., Beauregard, R., Katz, V. (1995). Linear Algebra, vol. 53.

Freestone, D.R., Karoly, P.J., Cook, M.J. (2017). A forward-looking review of seizure prediction. *Current Opinion in Neurology*, *30*(2), 167–173.

Gadhoumi, K., Gotman, J., Lina, J.M. (2015). Scale invariance properties of intracerebral EEG improve seizure prediction in mesial temporal lobe epilepsy. *PloS One*, *10*(4), e0121,182.

Gadhoumi, K., Lina, J.M., Mormann, F., Gotman, J. (2016). Seizure prediction for therapeutic devices: a review. *Journal of Neuroscience Methods*, *260*, 270–282.

Gantmacher, F.R. (1960). *Theory of matrices. 2V*. New York: Chelsea.

Ghaderyan, P., Abbasi, A., Sedaaghi, M.H. (2014). An efficient seizure prediction method using knn-based undersampling and linear frequency measures. *Journal of Neuroscience Methods*, *232*, 134–142.

Haemers, W.H., & Spence, E. (2004). Enumeration of cospectral graphs. *European Journal of Combinatorics*, *25*(2), 199–211.

Henry, B., Lovell, N., Camacho, F. (2012). Nonlinear dynamics time series analysis. *Nonlinear Biomedical Signal Processing: Dynamic Analysis and Modeling*, *2*, 1–39.

Hively, L. (2009). Prognostication of helicopter failure. ORNL/TM-2009, vol. 244.

Hively, L.M., & Ng, E.G. (1998). Integrated method for chaotic time series analysis. US Patent 5,815,413.

Hively, L., Clapp, N., Daw, C., Lawkins, W., Eisenstadt, M. (1995). Nonlinear analysis of EEG for epileptic seizures. ORNL/TM-12961, Oak Ridge National Laboratory, Oak Ridge.

Hively, L.M., Protopopescu, V.A., Munro, N.B. (2005). Enhancements in epilepsy forewarning via phase-space dissimilarity. *Journal of Clinical Neurophysiology*, *22*(6), 402–409.

Hively, L.M., McDonald, J.T., Munro, N., Cornelius, E. (2013). Forewarning of epileptic events from scalp EEG. In *Biomedical sciences and engineering conference (BSEC), 2013* (pp. 1–4): IEEE.

Huang, X., Altahat, S., Tran, D., Sharma, D. (2012). Human identification with electroencephalogram (eeg) signal processing. In *2012 International symposium on communications and information technologies (ISCIT)* (pp. 1021–1026). IEEE.

Ibrahim, S.W., Djemal, R., Alsuwailem, A., Gannouni, S. (2017). Electroencephalography (eeg)-based epileptic seizure prediction using entropy and k-nearest neighbor (knn). *Communications in Science and Technology*, *2*(1), 6–10.

Kandel, E.R., Schwartz, J.H., Jessell, T.M., Siegelbaum, S.A., Hudspeth, A.J. (2000). *Principles of neural science* (Vol. 4). New York: McGraw-Hill.

Kannathal, N., Min, L., Acharya, U., Sadasivan, P. (2006). Erratum: Entropies for detection of epilepsy in EEG (computer methods and programs in biomedicine (2005) 80(187–194). https://doi.org/10.1016/j.cmpb.2005.06.012.

Kantz, H., & Schreiber, T. (2004). *Nonlinear time series analysis* (Vol. 7). Cambridge: Cambridge University Press.

Karoly, P.J., Freestone, D.R., Boston, R., Grayden, D.B., Himes, D., Leyde, K., Seneviratne, U., Berkovic, S., O'Brien, T., Cook, M.J. (2016). Interictal spikes and epileptic seizures: their relationship and underlying rhythmicity. *Brain*, *139*(4), 1066–1078.

Kiral-Kornek, I., Roy, S., Nurse, E., Mashford, B., Karoly, P., Carroll, T., Payne, D., Saha, S., Baldassano, S., O'Brien, T., et al. (2017). Epileptic seizure prediction using big data and deep learning: toward a mobile system. *EBioMedicine*, *27*, 103–111.

Li, S., Zhou, W., Yuan, Q., Liu, Y. (2013). Seizure prediction using spike rate of intracranial EEG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *21*(6), 880–886.

Litt, B., Esteller, R., Echauz, J., D'Alessandro, M., Shor, R., Henry, T., Pennell, P., Epstein, C., Bakay, R., Dichter, M., et al. (2001). Epileptic seizures may begin hours in advance of clinical onset: a report of five patients. *Neuron*, *30*(1), 51–64.

Luckett, P., McDonald, J.T., Hively, L.M. (2017). Dissimilarity of graph invariant features from EEG phase-space analysis. *Computer Engineering and Information Technology*, *6*(3).

Martis, R.J., Acharya, U.R., Tan, J.H., Petznick, A., Tong, L., Chua, C.K., Ng, E.Y.K. (2013). Application of intrinsic time-scale decomposition (itd) to EEG signals for automated seizure prediction. *International Journal of Neural Systems*, *23*(5), 1350,023.

Meghdadi, A.H., Fazel-Rezai, R., Aghakhani, Y. (2017). Seizure prediction by nonlinear smoothness analysis of scalp eeg recording. *CMBES Proceedings*, *30*(1).

Mormann, F., Elger, C.E., Lehnertz, K. (2006). Seizure anticipation: from algorithms to clinical practice. *Current Opinion in Neurology*, *19*(2), 187–193.

Mormann, F., Andrzejak, R.G., Elger, C.E., Lehnertz, K. (2007). Seizure prediction: the long and winding road. *Brain*, *130*(2), 314–333.

Namazi, H., Kulish, V.V., Hussaini, J., Hussaini, J., Delaviz, A., Delaviz, F., Habibi, S., Ramezanpoor, S. (2016). A signal processing based analysis and prediction of seizure onset in patients with epilepsy. *Oncotarget*, *7*(1), 342.

Nan, X., & Jinghua, X. (1988). The fractal dimension of EEG as a physical measure of conscious human brain activities. *Bulletin of Mathematical Biology*, *50*(5), 559–565.

Osorio, I., Frei, M.G., Sornette, D., Milton, J. (2009). Pharmaco-resistant seizures: self-triggering capacity, scale-free properties and predictability? *European Journal of Neuroscience*, *30*(8), 1554–1558.

Osorio, I., Zaveri, H.P., Frei, M.G., Arthurs, S. (2016). *Epilepsy: the intersection of neurosciences, biology, mathematics, engineering, and physics*. Boca Raton: CRC Press.

Pauletti, A., Terrone, G., Shekh-Ahmad, T., Salamone, A., Ravizza, T., Rizzi, M., Pastore, A., Pascente, R., Liang, L.P., Villa, B.R., et al. (2017). Targeting oxidative stress improves disease outcomes in a rat model of acquired epilepsy. *Brain*, *140*(7), 1885–1899.

Sackellares, J.C. (2008). Seizure prediction. *Epilepsy Currents*, *8*(3), 55–59.

Sayama, H. (2015). Introduction to the modeling and analysis of complex systems. Open SUNY Textbooks.

Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980* (pp. 366–381). Berlin: Springer.

Truccolo, W., Donoghue, J.A., Hochberg, L.R., Eskandar, E.N., Madsen, J.R., Anderson, W.S., Brown, E.N., Halgren, E., Cash, S.S. (2011). Single-neuron dynamics in human focal epilepsy. *Nature Neuroscience*, *14*(5), 635–641.

Vahabi, Z., Amirfattahi, R., Shayegh, F., Ghassemi, F. (2015). Online epileptic seizure prediction using wavelet-based bi-phase correlation of electrical signals tomography. *International Journal of Neural Systems*, *25*(6), 1550,028.

Viglione, S., & Walsh, G. (1975). Proceedings: epileptic seizure prediction. *Electroencephalography and Clinical Neurophysiology*, *39*(4), 435.

Wang, S., Chaovalitwongse, W.A., Wong, S. (2013). Online seizure prediction using an adaptive learning approach. *IEEE Transactions on Knowledge and Data Engineering*, *25*(12), 2854–2866.

Williamson, J.R., Bliss, D.W., Browne, D.W., Narayanan, J.T. (2012). Seizure prediction using EEG spatiotemporal correlation structure. *Epilepsy & Behavior*, *25*(2), 230–238.

Wilson, R.C., & Zhu, P. (2008). A study of graph spectra for comparing graphs and trees. *Pattern Recognition*, *41*(9), 2833–2841.

World Health Organization (2014). World health organization. http://www.who.int/mediacentre/factsheets/fs999/en/. Website.

Xiao, C., Wang, S., Iasemidis, L., Wong, S., Chaovalitwongse, W.A. (2017). An adaptive pattern learning framework to personalize online seizure prediction. *IEEE Transactions on Big Data*, (1), 1-1.

Yang, Y., Wang, Y., Wu, Q.J., Lin, X., Liu, M. (2015). Progressive learning machine: a new approach for general hybrid system approximation. *IEEE Transactions on Neural Networks and Learning Systems*, *26*(9), 1855–1874.

Yoo, Y. (2017). On predicting epileptic seizures from intracranial electroencephalography. *Biomedical Engineering Letters*, *7*(1), 1–5.

Zandi, A.S., Tafreshi, R., Javidan, M., Dumont, G.A. (2010). Predicting temporal lobe epileptic seizures based on zero-crossing interval analysis in scalp EEG. In *2010 annual international conference of the IEEE engineering in medicine and biology society (EMBC)* (pp. 5537–5540). IEEE.

Zappasodi, F., Olejarczyk, E., Marzetti, L., Assenza, G., Pizzella, V., Tecchio, F. (2014). Fractal dimension of EEG activity senses neuronal impairment in acute stroke. *PLoS One*, *9*(6), e100,199.

Zheng, Y., Wang, G., Li, K., Bao, G., Wang, J. (2014). Epileptic seizure prediction using phase synchronization based on bivariate empirical mode decomposition. *Clinical Neurophysiology*, *125*(6), 1104–1111.

Zheng, Y., Zhang, H., Yu, Y. (2015). Detecting collective anomalies from multiple spatio-temporal datasets across different domains. In *Proceedings of the 23rd SIGSPATIAL international conference on advances in geographic information systems* (p. 2). ACM.

Zhu, J., He, W., Yang, H. (2008). Contrastive analysis of correlation dimension of EEG signals between normal and pathological groups. In *Automation congress, 2008. WAC 2008. World* (pp. 1–4). IEEE.